

PROMEMORIOR FRÅN P/STM

NR 6

METODER FÖR EVALVERING AV  
NOGGRANNHETEN I SCBs STATISTIK.  
EN ÖVERSIKT.

AV JÖRGEN DALÉN

## INLEDNING

### TILL

**Promemorior från P/STM / Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån, 1978-1986. – Nr 1-24.**

#### **Efterföljare:**

Promemorior från U/STM / Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån, 1986. – Nr 25-28.

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

Promemorior från P/STM 1981:6. Metoder för evalvering av noggrannheten i SCBs statistik. En översikt / Jörgen Dalén.  
Digitaliserad av Statistiska centralbyrån (SCB) 2016.

PROMEMORIOR FRÅN P/STM

NR 6

METODER FÖR EVALVERING AV  
NOGGRANNHETEN I SCBs STATISTIK.  
EN ÖVERSIKT.

AV JÖRGEN DALEN'



INNEHÅLLSFÖRTECKNING TILL  
METODER FÖR EVALVERING AV NOGGRANNHETEN I SCBs STATISTIK.  
EN ÖVERSIKT.

INLEDNING

- 1 Fel i olika typer av statistik
  - 1.1 Surveyen
  - 1.2 Fel i en survey
  - 1.3 Händelsebaserad statistik
  - 1.4 Fel i händelsebaserad statistik
  - 1.5 Registerbaserad statistik
  - 1.6 Fel i registerbaserad statistik
  - 1.7 Andra typer av statistik
  - 1.8 Sammanfattande kommentar
  
- 2 Felmodeller och felredovisning
  - 2.1 Medelkvadratavvikelsen
  - 2.2 Felprofilen
  - 2.3 Denna rapports systematik
  - 2.4 Felmått i evalveringsundersökningar
    - 2.4.1 Felmått på objektsnivå
  
- 3 Evalveringsmetoder för olika fel
  - 3.1 Övergripande evalveringsmetoder
    - 3.1.1 Jämförelse med aggregerade data från andra källor
    - 3.1.2 Konsistensstudier
    - 3.1.3 Preliminära-definitiva resultat
  - 3.2 Evalvering av fel på objektsnivå
    - 3.2.1 Jämförelser med annat material på objektsnivå
    - 3.2.2 Kontroll av uppgiftslämnande vid källan/återintervju med bättre mätmetod
    - 3.2.3 Indirekta kontroller av uppgiftslämnandet
    - 3.2.4 Replikering
    - 3.2.5 Evalvering av kodning
    - 3.2.6 Evalvering av stansning
    - 3.2.7 Evalvering av granskning/rättning
  - 3.3 Evalvering av fel på populationsnivå
    - 3.3.1 Jämförelser med sant värde
    - 3.3.2 Strukturjämförelser mellan svarande och icke-svarande
    - 3.3.3 Studier av variationer över svarsomgångar
    - 3.3.4 Jämförelser mellan skattningsmetoder
    - 3.3.5 Matchning av register
    - 3.3.6 Analys av eftersläpningseffekter i register
    - 3.3.7 Kontroll av ofullständig rapportering i händelsebaserad statistik
    - 3.3.8 Kontroll av bearbetningsfel på populationsnivå



## INLEDNING

Statistiska undersökningar vid SCB syftar till att i kvantitativa termer belysa någon aspekt av den samhällsliga verklighet, som vi har omkring oss. Resultaten från undersökningarna utgör underlag för politiska och ekonomiska beslut och fördjupad forskning samt bidrar till den totala kunskapsbas, som ligger till grund för t ex utbildning och demokratisk opinionsbildning.

För att statistiken skall fylla denna funktion krävs att dess kvalitet är känd samt uppfyller de minimikrav, som dess syfte(-n) ställer.

Ett stort arbete läggs normalt ner i SCBs undersökningar på att minimera de slutliga felen. Det kan gälla påminnelser till uppgiftslämnare eller gransknings- och rättningsprogram. Mer sällsynt är att arbete läggs ner på att analysera kvaliteten i den slutligen framtagna produkten - på att evalvera statistiken.

Som motivering för denna resursfördelning brukar ibland hävdas att "det är bättre att använda resurserna till att rätta till felen i statistiken än att redovisa dem". Vi vill opponera oss mot detta synsätt, av flera skäl. Men låt oss först säga något om vad vi menar med evalvering.

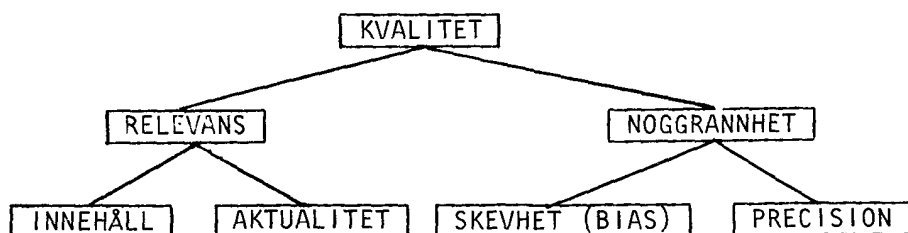
En evalvering - utvärdering - av en statistisk undersökning syftar till att ge kunskap om undersökningens kvalitet. Denna kunskap kan användas till två saker. Dels till att upplysa användarna om den osäkerhet som vidlåder de redovisade statistiska resultaten - kvalitetsredovisning. Dels till att göra det möjligt för producenten att på ett rationellt sätt planera den statistiska produktionsprocessen och besluta om hur resurserna optimalt bör fördelas mellan olika produktionsmoment i syfte att maximera den utgående kvaliteten vid given kostnad eller att minimera kostnaden för en given kvalitetsnivå.

Av detta följer att en evalvering av en statistikprodukt har ett egenvärde i den meningen att den klargör statistikens användbarhet. För endast en evalverad och kvalitetsredovisad statistik gör det möjligt att avgöra huruvida statistiken är användbar för ett visst syfte. Dessutom tillkommer att en rationell fördelning av resurserna mellan olika produktionsmoment förutsätter en - åtminstone grov - kännedom om förhållandet mellan felriskerna i de respektive momenten.

Därför är en evalvering av en statistikprodukt alltid motiverad. Problemet är snarast att bestämma ambitionsnivån för en sådan evalvering. Denna sammanhänger naturligtvis dels med kostnaderna för evalveringen, dels med vad man förväntar sig få ut av den i form av ökad kunskap om felsituationen. Allmänt sett bör målet vara att felsituationen bringas under kontroll, dvs att statistikproducenten får en sådan kunskap om statistikens kvalitet att hon med eftertryck kan uttala sig om de slutsatser som de statistiska resultaten berättigar till att dra respektive kan varna för en sådan användning av uppgifterna, som deras kvalitet inte håller för.

En statistikproducent bör också kunna redovisa en bedömning av de totala felmarginaler en användare skall behöva kalylera med, när han ska använda resultaten.

Innan vi går vidare bör det kvalitetsbegrepp vi använder närmare preciseras. Inom SCB brukar det delas upp enligt följande schema:





Evalvering av en undersöknings relevans ingår som en del av vad som vid SCB går under benämningen behovsanalys. Det innebär att kartlägga hur den färdiga statistiken används, huruvida de planerade syftena med statistiken kan tillgodoses. Hur påverkas användningen av definitionen av population och variabler, av aktualitet m m? Vilka nya användningar skulle kunna tillkomma om ändringar i dessa avseenden gjordes?

Evalvering av relevansen är ett mycket viktigt område i statistikproduktionen. Vi skall dock inte gå in närmare på det i denna rapport, eftersom det kräver helt andra arbetsmetoder än en noggrannhetsorienterad evalvering, metoder som ofta är specifika för ett visst ämnesområde.

Denna rapport skall alltså begränsa sig till evalvering av noggrannheten - ett i sig tillräckligt stort ämne.

Rapporten inleds med en genomgång av felsituationen i några olika typer av statistiska undersökningar som finns vid SCB. Därefter görs en strukturering av felen i vissa huvudkategorier. Till sist presenteras ett antal tekniker för att evalvera olika typer av fel och referenser görs till evalveringsundersökningar, där de respektive metoderna kommit till användning.

Arbetet med denna rapport har genomförts inom projektet KAPLAN - planering av kvalitetskontrollarbete vid SCB. Förutom författaren har i olika skeden av arbetet Sven Näsman och Jan Håkan Wretman medverkat. Värdefulla synpunkter har lämnats av Ingrid Lyberg, Lars Lyberg och andra inom P/STM och inom styrgruppen för KAPLAN.

## 1 Fel i olika typer av statistik

Låt oss betrakta några olika typer av statistik, som det ur metodologisk synvinkel är fruktbart att särskilja. Fel-situationen skiljer sig på ett kvalitativt sätt mellan dessa och därför måste också evalveringsmetoderna bli delvis annorlunda. Tre olika slag av statistik skall särskilt skärskådas, nämligen surveyen, händelsebaserad statistik och registerbaserad statistik.

### 1.1 Surveyen

I statistisk litteratur används det engelska begreppet survey för att beskriva en typ av undersökning, där man samlar in data för att dra slutsatser om aggregat av enheter, populationer. Dalenius (1974) har gett en strikt och något snävare definition av vad en survey är, som vi refererar här i avkortad form:

- (1) Surveyen gäller en väl definierad mängd objekt, en *population*.
- (2) *Populationen* har en eller flera mätbara egenskaper.
- (3) Man vill beskriva *populationen* med parametrar som definieras i termer av dessa egenskaper.
- (4) För att få observationskontakt med *populationen* behöver vi en ram, som operationellt representerar *populationen*.
- (5) Ett urval av objekt väljs från *ramen* med hjälp av en sannolikhetsmekanism (ett specialfall är att alla objekt väljs, vilket innebär en totalundersökning).
- (6) Observationer görs av *urvalet* med hjälp av en viss mätmetod.
- (7) Ett estimationsförfarande appliceras, vilket gör det möjligt att skatta *populationens* värden.

En survey är vid SCB typiskt sett en undersökning byggd på ett urval av individer eller hushåll ur RTB eller av företag ur CFR. Åtskilliga andra varianter förekommer naturligtvis. I en survey är alltså det centrala att studera någon egenskap hos de objekt som tillhör populationen.

Referens:

Dalenius T: Ends and means of total survey design, Forskningsprojektet FEL I UNDERSÖKNINGAR, Stockholms universitet, Statistiska institutionen, Juni 1974, sid 2-12.

1.2 Fel i en survey

Det första steget i en survey är normalt att på grundval av ett existerande sakproblem fastställa en lämplig undersökningspopulation avgränsad i tid och rum. Därefter måste man välja en ram där populationen finns representerad. Ofta finns nu en skillnad mellan undersökningspopulationen och denna ram, som kan ge upphov till ett täckningsfel. Täckningsfel kan bero på övertäckning (objekt i ramen ingår ej i populationen) eller undertäckning (objekt i populationen ingår ej i ramen).

Från ramen dras (ibland) ett urval av objekt, vilket ger upphov till ett urvalsfel som utgörs av skillnaden mellan den skattning som erhålls från urvalet och den som skulle ha erhållits från hela populationen.

Det är ofta inte möjligt att erhålla uppgifter från alla objekt som ingår i undersökningen. Detta resulterar i ett bortfallsfel som utgörs av skillnaden mellan den skattning som erhålls från de svarande och den som skulle erhållits från hela urvalet (populationen vid totalundersökning).

De erhållna uppgifterna är inte alltid korrekta. Medvetet eller omedvetet kan en uppgiftslämnare lämna felaktiga svar t ex på grund av bristfälliga anvisningar. Eller ett mätinstrument kan ge ett felaktigt värde. Dessa fel brukar kallas mätfel (ibland också uppgiftslämnarfel eller observationsfel).

I intervjuundersökningar uppkommer intervjuarfel p g a intervjuarnas olika sätt att ställa frågor och tolka svar. Detta kan betraktas som en särskild typ av mätfel.

Undersökningens sista skede är så bearbetningen vid SCB. Blanketter kodas, granskas, dataregistreras, tabeller programmeras och framställs och resultat skrivs ut och publiceras. Alla dessa produktionsled vid SCB kan ge upphov till bearbetningsfel, som naturligtvis kan uppdelas ytterligare i kodningsfel, granskningsfel, dataregistreringsfel, programmeringsfel och utskriftsfel.

Det är inte ovanligt att mätfelet dominerar felbilden i en survey även där urvalsfelet är ganska stort. Till skillnad från urvalsfelet går det heller inte att skatta utan vidare. Det är därför naturligt att många evalveringsundersökningar inriktar sig just på mätfelet.

### 1.3 Händelsebaserad statistik

En händelsebaserad statistik karaktäriseras av att en viss händelse ger upphov till en statistikrapport, t ex en blankett som skickas in till SCB. Till skillnad från vad som är fallet vid en survey, kommer alltså inte initiativet till statistikrapporteringen i de enskilda fallen från SCB.

Populationen utgörs av mängden av händelser som ägt rum under en viss tidsperiod.

Exempel på sådan statistik är

- . utrikeshandelsstatistiken, som bygger på tullsedlar som ifylls varje gång en vara passerar gränsen
- . brottsstatistiken, där varje brottsanmälan ger upphov till en statistikrapport.

I denna typ av statistik är huvudintresset oftast knutet till att mäta populationens storlek eller någon funktion av denna. Beskrivningen av populationens egenskaper kan också vara

viktig men kommer i andra hand. Eftersom populationen är okänd saknas en ram för denna.

Av surveykriteriets sju punkter är sålunda (3) - (5) ej helt uppfyllda och som vi ska se nedan så passar inte hela den för surveysituationen uppbyggda begreppsapparaten vad gäller feltyper för denna typ av statistik.

#### 1.4 Fel i händelsebaserad statistik

Den helt avgörande noggrannhetsfaktorn i en händelsebaserad statistik är hur väl man lyckas täcka populationen, dvs hur fullständig rapportering av händelserna man uppnår.

Motsvarande fel kan alternativt betecknas undertäckning eller bortfall; någon åtskillnad mellan dessa begrepp blir ej meningsfull att göra. (I brottsstatistiken används begreppet mörkertal för den del av brottsligheten som aldrig anmäls och som därför inte kommer med i undersökningen.) Här gäller också att undertäcknings-/bortfallsfelet är identiskt lika stort som undertäckningen/bortfallet självt. I en survey däremot beror detta fel på undertäckningens respektive bortfallets snedvridande effekt.

En alternativ ansats är att betrakta uppgiftslämnarna som undersökningspopulation och antalet händelser som en mätvariabel. Felsituationen kommer då att bli mer lik den i en survey. I de fall uppgiftslämnarna är myndigheter, institutioner etc blir då bortfalls- och täckningsfelen i de flesta fall noll.

Normalt bygger den händelsebaserade statistiken på en totalundersökning, varför något urvalsfel inte förekommer. Mätfel och bearbetningsfel förekommer i princip i samma utsträckning som i surveyundersökningar.

## 1.5 Registerbaserad statistik

Vi betraktar här en population med följande karaktäristika: Ett objekt föds, varvid det registreras. Under objektets livstid sker förändringar i dess egenskaper, varvid ändringar förs in i registret. Slutligen dör objektet, varvid det avförs ur registret. En statistik som direkt bygger på denna form av register kallar vi en registerbaserad statistik.

Karaktäristiskt för registret är att målsättningen är att det vid vissa tidpunkter skall omfatta hela populationen. Detta förutsätter ett utarbetat uppdateringsystem.

Inom SCB finns åtskilliga exempel på registerbaserad statistik:

- . Den löpande befolkningsstatistiken, som bygger på RTB
- . Företagsredovisning ur CFR
- . Lärarstatistik ur lärarregistret

Även denna typ av statistik skiljer sig avsevärt från survey-situationen. Huvudmålsättningen är liksom i den händelsebaserade statistiken oftast att mäta populationens storlek. Skillnaden mellan de två är att registret har en tidsdimension; i detta ingår objekt som existerar under en viss tidsperiod. Den händelsebaserade statistiken däremot bygger på objekt som inte har någon tidsdimension utan är definierade för en viss tidpunkt.

## 1.6 Fel i registerbaserad statistik

Populationen i ett register kan avvika från undersökningspopulationen av flera orsaker:

- . Avgränsningen av de objekt som skall ingå i registerpopulationen är ibland svår att strikt upprätthålla, vilket ger upphov till såväl över- som undertäckningsfel.

- . Vissa objekt registreras aldrig. Detta ger upphov till ofullständigheter i registret som alternativt kan benämnas undertäckning eller bortfall.
- . Ett objekt registreras vid en tidpunkt skild från dess verkliga födelse-tidpunkt (normalt för sent). Detta ger vid varje enskild tidpunkt upphov till ett täckningsfel (normalt undertäckning).
- . Ändringar i ett objekts egenskaper registreras inte alls eller vid fel tidpunkt, vilket ger upphov till ett mätfel i statistik från registret.
- . Ett objekts död registreras inte alls eller vid fel tidpunkt (normalt för sent). Detta ger upphov till en övertäckning i statistiken.

Därutöver finns samma mätfel och bearbetningsfel som i andra typer av undersökningar. Liksom den händelsebaserade statistiken inriktar sig den registerbaserade statistiken vanligen på hela populationen, varför något urvalsfel inte förekommer.

### 1.7 Andra typer av statistik

Vid SCB förekommer även annan statistikproduktion än den som faller in under kategorin 1.1-1.6. Här skall några sådana typer i korthet nämnas.

Olika typer av index kan egentligen komma in under vilken som helst av ovanstående rubriker. Index är ju bara ett alternativt sätt att redovisa statistiska resultat, vilket fokuserar förändringar i tiden. Prisindex är dock en speciell typ av statistik med en komplicerad felsituation.

Statistiska hybridprodukter är t ex nationalräkenskaperna och finansräkenskaperna. Det utgör en sammanställning och bearbetning av annat statistiskt material. Felsituationen i denna typ av statistik är synnerligen komplex.

Prognoser är ytterligare en form av undersökning, som görs vid SCB. Det dominerande felet i en prognos härrör normalt från prognosmodellen som sådan och dess avvikelse från den påföljande verkligheten.

Vi ska inte här behandla evalvering av dessa typer av statistik utan begränsa oss till de i 1.1-1.6 nämnda typerna.

## 1.8 Sammanfattande kommentar

Varje statistisk undersökning har på sitt sätt en unik fel-situation. Speciella typer av fel uppkommer, som inte självklart kan åsättas en känd etikett.

Därför måste det första steget mot en evalvering av noggrannheten vara en genomgång och strukturering av de tänkbara felkällorna i undersökningen. Ofta har en statistikproducent sådana kunskaper att han från början har en uppfattning om vilka felkällor, som är av betydelse och vilka som är försumbara.



## 2 Felmodeller och felredovisning

Utifrån kunskap om de enskilda felkällorna i statistiken är det inte möjligt att göra något uttalande om det totala felets storlek. Men det är detta som en konsumentorienterad evalvering syftar till. Därför finns ett behov av en felmodell, som på något sätt kan härleda ett mått på totalfelet utifrån de enskilda felkällorna.

En felmodell är även nödvändig för att kunna analysera de resultat man får ut ur en evalveringsundersökning. Man måste kunna skilja slumpmässiga fel från systematiska och se att ett konstaterat fel kan vara en samlad effekt av flera olika felkällor.

Vi ska här redogöra för ett par försök att skapa felmodeller eller andra typer av strukturer över felen. Därefter redogörs för den systematik som kommer att ligga till grund för den fortsatta framställningen.

### 2.1 Medelkvadratavvikelsen

Den mest utarbetade modellen för fel i statistiska undersökningar utgår från begreppet medelkvadratavvikelsen, som är ett mått på den totala kvaliteten i en undersökning i vilken även relevansen brukar inräknas.

Man börjar med att formulera ett idealt undersökningsmål ( $z$ ). Det ideala målet skulle uppnås, om man kunde tillgodose informationsbehovet fullständigt.

Men i praktiken är det ideala målet ofta ouppnåeligt. Med kunskap om tillgängliga metodologiska resurser måste därför det ideala målet revideras och preciseras, så att det blir operationellt meningsfullt och kan ligga till grund för planeringen av en undersökning. Denna process leder till ett definierat undersökningsmål ( $x$ ), som alltså blir utgångspunkt för undersökningsplaneringen.

När undersökningen sedan genomförs, resulterar den i ett observerat surveyresultat ( $Y$ ), som av många anledningar inte är exakt lika med det definierade målet ( $x$ ). Vi talar också om ett förväntat surveyresultat ( $y$ ), vilket uppfattas som det värde man i genomsnitt skulle få på  $Y$  vid en tänkt följd av oändligt många upprepningar av samma undersökning under lika betingelser.

En surveys kvalitet har att göra med hur bra man lyckas träffa målet. Vi tänker oss i fortsättningen att  $z$ ,  $x$  och  $y$  är okända reella tal, medan  $Y$  är en reell stokastisk variabel, vilken uppfattas som en skattning av  $z$  (eller av  $x$ , beroende på hur man ser det). Det totala skattningsfelet m a p det ideala målet blir  $Y - z$ , vilket kan delas upp i komponenterna

$$\begin{aligned} Y - z &= (Y - y) + (y - x) + (x - z) \\ &= \text{slumpfel} + \text{bias} + \text{relevansfel} \end{aligned}$$

Som mått på en skattnings osäkerhet väljer vi att använda medelkvadratavvikelsen m a p det ideala målet,  $E(Y-z)^2$ , vilket kan delas upp i komponenterna

$$\begin{aligned} E(Y-z)^2 &= E(Y-y)^2 + (y-x)^2 + (x-z)^2 + 2(y-x)(x-z) \\ &= \text{varians} + (\text{bias})^2 + (\text{relevansfel})^2 + \\ &\quad + 2(\text{bias}) \times (\text{relevansfel}) \end{aligned}$$

I denna framställning är vi primärt intresserade av att mäta det totala felet. Man kan då nöja sig med att tala om det totala skattningsfelet m a p det definierade målet och medelkvadratavvikelsen m a p det definierade målet:

$$Y-x = (Y-y) + (y-x) = \text{slumpfel} + \text{bias}$$

$$E(Y-x)^2 = E(Y-y)^2 + (y-x)^2 = \text{varians} + (\text{bias})^2$$

På denna grundval kan ytterligare mer förfinade modeller byggas upp, där variansen delas upp i olika komponenter. En sådan modell är US' Bureau of the Census' felmodell.

Referenser:

Hansen M H, Hurwitz W N and Pritzker L: The estimation and interpretation of gross differences and the simple response variance. Contributions to statistics, presented to professor P C Mahalanobis on the occasion of his 70th birthday, Pergamon Press, Oxford, England, 1964.

Hansen M H, Hurwitz W N and Pritzker L: Standardization of procedures for the evaluation of data: Measurement errors and statistical standards in the Bureau of the Census. Bulletin of the International Statistical Institute, Vol 42, Book 1, 49-66.

Wretman J: Beskrivning av en surveyfelmodell, SCB P/STM, Stencil, 1978-08-22.

## 2.2 Felprofilen

Medelkvadratavvikelseansatsen har inte kommit till användning i så stor utsträckning, som dess skapare kanske hade hoppats på. Detta sammanhänger med att dess komponenter inte entydigt kan hänföras till en enskild felkälla. Komponenterna blir därigenom nästan lika svåra att uppskatta som medelkvadratavvikelsen själv. Det är också dyrbart att genomföra de experiment m m som ofta krävs för att korrekt kunna skatta dessa komponenter. Dessutom har inte bortfallsfelen och täckningsfelen kunnat inkorporeras i medelkvadratavvikelsen.

På senare år har därför en annan ansats föreslagits, den s k felprofilen.

En felprofil är en systematisk och omfattande redovisning av alla produktionsmoment i en statistisk undersökning, som ger ett bidrag till det totala felet (Dalenius, 1977). Syftet med denna kan vara att göra en teknisk dokumentation av undersökningen för producenten själv och andra som har anledning att sätta sig in i detalj i undersökningens framställning. Vissa förespråkare för felprofilansatsen vill också betona ett användarorienterat syfte med denna, men mot detta kan invändas att en användare av statistik inte har behov av en omfattande teknisk diskussion av sin siffra utan av ett totalmått på dess tillförlitlighet.

Som slutmål för en producentorienterad evalvering och disposition av en teknisk rapport kan dock felprofilen vara lämplig. Man kan spalta upp en undersökning i sina beståndsdelar och studera felbidragen från var och en av dessa. Dalenius ger följande exempel på en sådan uppdelning, som är tillämplig i första hand på surveyundersökningar:

- 1 Undersökningens planering
  - 1.1 Val av undersökningsvariabler
  - 1.2 Val av undersökningspopulation
- 2 Undersökningens utförande
  - 2.1 Skapa observationskontakt med populationen: utveckla ramen, välja urvalet etc
  - 2.2 Samla in data
  - 2.3 Bearbeta data
  - 2.4 Beräkna de statistiska resultaten
  - 2.5 Beräkna mått på felen

I denna uppställning betraktas även relevans-"felet" i punkt 1, vilket skiljer sig från den ansats vi i övrigt valt i denna framställning.

#### Referenser:

Brooks, C A and Baïlar, B A (1977): An error profile: employment as measured by the Current Population Survey. American Statistical Association, Proceedings of the Social Statistics Section 1977 Part 1, Washington D.C. 1978, 26-34.

Dalenius T (1977): Strain at a gnat and swallow a camel: or, the problem of measuring sampling and non-sampling errors. American Statistical Association, Proceedings of the Social Statistics Section 1977 Part 1, Washington D. C. 1978, 21-25

Madow, L H (1977): An error profile: employment as measured by the current employment statistics program. American Statistical Association, Proceedings of the Social Statistics Section 1977 Part 1, Washington D.C. 1978, 35-44.

### 2.3 Denna rapporters systematik

I denna rapport har en annan ansats valts, som har förefallit lämplig i syfte att skilja ut tillämpningsområdena för olika evalveringsmetoder. Huvudpoängen i denna ansats är att skilja fel som uppstår i mätvärdena för enskilda undersökningsobjekt (fel på objektsnivå) från fel som uppstår i inferens-/uppräkningsituationen (fel på populationsnivå). Tillsammans med avgränsningen mellan slumpmässiga och systematiska fel kan man då definiera följande fyra klasser av fel:

I Slumpmässiga fel på objektsnivå. Felaktigheter i mät-, insamlings- och bearbetningsförfarandet som är av slumpmässig karaktär och som har ett förväntat värde lika med noll sett över alla objekt.

II Systematiska fel på objektsnivå. Felaktigheter i mät-, insamlings- och bearbetningsförfarande, som systematiskt snedvrider de enskilda mätvärdena i en bestämd riktning.

III Slumpmässiga fel på populationsnivå. Detta är främst fråga om urvalsfel, som genereras p g a att endast en del av undersökningspopulationen studeras. Men även bortfallet och överresp undertäckningen ger ett bidrag till det slumpmässiga felet.

IV Systematiska fel på populationsnivå. Till denna grupp hör först och främst bortfallets och undertäckningens snedvriddande effekt men även andra typer av bias i uppräkningsförfarandet.

I den följande sammanställningen kommer vi för varje evalveringsmetod att redovisa vilken/vilka av dessa klasser av fel som de lämpar sig för att studera.

### 2.4 Felmått i evalveringsundersökningar

Olika slags fel måste mätas med olika mått. Generellt brukar variansen (standardavvikelsen) vara ett lämpligt mått på det slumpmässiga felet medan bias brukar få beteckna det systematiska felet.

Ur variansen kan under vissa förutsättningar ett osäkerhets-  
tal härledas, som svarar mot ett visst konfidensintervall  
kring det skattade värdet ( $2 \times \sqrt{V}$  leder till ett 95 %  
konfidensintervall). Detta konfidensintervall motsvarar  
undersökningens urvalsfel. När en undersökning är behäftad  
med slumpmässiga mätfel kommer en på vanligt sätt beräknad  
varians att även inkludera dessa (förutsatt att urvalsfrak-  
tionen är liten och att felen för olika objekt inte är korre-  
lerade).

Bortfalls- och täckningsfel leder till en bias i den slutliga  
skattningen, som i vissa fall kan uppskattas. Bidrag till  
biasen i undersökningen kan dock komma från alla felkällor i  
undersökningen.

#### 2.4.1 Felmått på objektsnivå

När ett sant värde beräknats för (åtminstone ett sannolik-  
hetsurval av) objekten blir det aktuellt att använda ett  
par speciella felmått. Dessa har olika syften och det är  
väsentligt att använda det för det aktuella syftet mest  
relevanta.

Allmänt kan vi betrakta problemet på följande sätt:

Låt  $x_i$  vara värdet av en viss variabel för objekt  $i$  i huvud-  
undersökningen. Låt  $y_i$  vara (det samma) värdet av samma  
variabel för samma objekt  $i$  i evalveringsundersökningen.

Om evalveringsundersökningen är en totalundersökning blir nu  
nettofelet för variabeln

$$\sum_{i=1}^N (x_i - y_i)$$

där  $N$  är antalet objekt  $i$  i den aktuella redovisningsgruppen.  
Man kan också tala om det relativa nettofelet eller netto-  
felet i procent. Detta är

$$\frac{\sum_{i=1}^N (x_i - y_i) \cdot 100}{\sum_{i=1}^N x_i} \%$$

Nettotofelet är ett uttryck för felens effekt på de redovisade resultaten och är därför det som primärt intresserar en statistikanvändare.

Bruttotofelet för samma variabel kan t ex definieras som

$$\sum_{i=1}^N |x_i - y_i|,$$

dvs summan av felens absolutbelopp. (Andra beräkningssätt är naturligtvis också tänkbara, t ex roten ur kvadratsumman.)

Det relativa bruttotofelet kan då uttryckas som

$$\frac{\sum_{i=1}^N |x_i - y_i|}{\sum_{i=1}^N x_i}$$

Bruttotofelet är ett i första hand producentorienterat felmått, som ger en upplysning om storleken på "felen i sig". Det visar kvaliteten på ett visst eller vissa produktionsmoment (t ex datainsamling eller kodning). I de fall man beräknar bruttotofel för en övergripande redovisningsnivå (t ex hela riket) så kan emellertid detta felmått ge en fingervisning om felriskerna på lägre redovisningsnivåer (t ex kommuner). På detta sätt kan även bruttotofelet ha ett användarorienterat syfte.

Ett tredje felmått skulle kunna vara antalet objekt med fel resp andelen objekt med fel som motsvarande relativa felmått. Även detta felmått är av bruttokaraktär och har ett liknande syfte som det vanliga bruttotofelet. Skillnaden är att det senare tar hänsyn till felens storlek, vilket i allmänhet är en fördel.

Låt oss belysa de olika felmåttan med hjälp av följande exempel. Vi har här värden från såväl den ordinarie undersökningen som evalveringsundersökningen för en redovisningsgrupp som omfattar tio objekt.

Objekt nr	Värde erhållet vid		Fel
	Evalvering	Ord. undersökn.	
1	150	175	25
2	200	200	0
3	0	100	100
4	300	300	0
5	400	400	0
6	250	100	-150
7	100	200	100
8	100	100	0
9	50	50	0
10	100	100	0
Summa	1 650	1 725	

Nettofelet blir här  $1\ 725 - 1\ 650 = 75$  och det relativa nettofelet  $75 \times 100/1\ 725 \% = 4,3 \%$ .

Bruttofelet blir  $(25 + 100 + 150 + 100) = 375$  och det relativa bruttofelet  $375 \times 100/1\ 725 = 22 \%$

Antalet objekt med fel är fyra och motsvarande andel blir följaktligen  $40 \%$ .

Om evalveringsundersökningen som i de flesta fall bygger på ett sannolikhetsurval måste naturligtvis felskattningarna förses med en uppräkningsfaktor. Skattningarna blir då också utsatta för slumpmässiga variationer till följd av urvals-förfarandet.

Det i exemplet ovan redovisade nettofelet bygger på den implicita förutsättningen att man syftar till att skatta en totalsumma  $-\sum_{i=1}^N y_i$ . Om syftet i stället vore att göra en klassindelning av materialet skulle vi t ex kunna få någon av följande situationer.



Tablå A

Antal objekt	Värde			
	0	1-100	101-200	201-
enl evalv	1	4	2	3
enl ord. und.	0	5	3	2
Rel nettofel	∞	20	33	-50

Tablå B

Antal objekt	Värde		
	-100	101-250	251-
enl evalv	5	3	2
enl ord. und.	5	3	2
Rel nettofel	0	0	0

Vi ser här att nettofelet blir helt beroende av de valda redovisningsgrupperna. Väljer man klassgränserna på ett visst sätt kan felen bli av en helt annan storleksordning än eljest. (Att skillnaderna blir så stora som mellan tablå A och B ovan beror naturligtvis också på det lilla antalet objekt i vårt exempel.)

I Henkel - Lyberg (1979) redogörs utförligare för dessa felmått i fallet med klassindelad material.

#### Referens:

Henkel M, Lyberg I: Några kvalitetsmått i samband med evalveringsundersökningar. SCB P/STM RIS, Stencil 1979-10-01

### 3 Evalveringsmetoder för olika fel

I detta avsnitt presenteras en uppsättning evalveringsmetoder. De som valts ut bedöms vara tillämpbara någorlunda ofta på SCBs statistik.

Metoderna presenteras ganska kortfattat och utan någon genomgång av den matematik, som erfordras för analys av resultaten. Denna finns i stället redogjord för i vissa av de referenser som anges efter varje avsnitt.

#### 3.1 Övergripande evalveringsmetoder

Att uppskatta totalfelet är den centrala målsättningen för varje konsumentorienterad evalveringsundersökning. Någon metod som direkt leder till en objektiv skattning av totalfelet existerar dock inte.

Nedan presenteras några enkla ansatser som syftar till att ge en indikation på det sammanlagda fel, som kan förekomma i en undersökning.

På grund av sin enkelhet är metoderna ofta lämpliga som en inledning till ett större evalveringsarbete. Genom att ge en indikation på felens troliga omfattning kan de också utgöra ett underlag för beslut om hur stora resurser som bör satsas på evalveringsarbetet.

##### 3.1.1 Jämförelse med aggregerade data från andra källor

Ofta produceras statistik inom ett visst område på flera håll inom och utom SCB. I andra fall kan delgrupper inom en viss undersökning överensstämma med data från andra källor.

Exempel på sådana situationer är:

- uppgifter om sysselsättning från AKU, FoB och företagsstatistik

- . uppgifter om arbetslöshet från SCB och AMS
- . uppgifter om den privata konsumtionen från hushållsbudgetundersökningen och från nationalräkenskaperna.

I dessa fall kan en jämförelse göras och skillnaden vara en indikation på osäkerheten i resultaten.

Det finns emellertid åtskilliga problem med en sådan jämförelse. För det första finns ofta periodiserings- eller andra definitionsskillnader mellan materialen. Om dessa är stora blir jämförelsen meningslös.

För det andra är tolkningen av de konstaterade skillnaderna inte självklar. Endast om den källa man jämför med kan bedömas ha en överlägsen noggrannhet, kan skillnaden tolkas som ett mått på totalfelet. I de fall jämförelsekällans noggrannhet är underlägsen blir jämförelsen meningslös. I övriga fall kan skillnaden göras till föremål för expertbedömningar som utnyttjar ytterligare information eller för fördjupade evalveringsundersökningar.

För det tredje måste man se upp med att två olika undersökningar kan ha gemensamma felkällor om de t ex baserar sig på ett gemensamt källmaterial eller om uppgiftslämnare har en benägenhet till systematiskt felsvarande på vissa typer av frågor, som olika undersökningar har ställt. Det totala felet kan då mycket väl vara större än den konstaterade skillnaden.

Ett exempel på hur denna metod kan användas utgör hushållsbudgetundersökningen 1978 (HBU). Där görs en jämförelse med nationalräkenskapernas (NR) uppgifter avseende den privata konsumtionen uppdelad på poster. Skillnaderna görs till föremål för bedömningar. Efter att effekter av definitionsskillnader rensats bort bedöms för olika poster den relativa noggrannheten hos de olika statistikuppgifterna.

Referenser:

SOS Hushållsbudgetundersökningen 1978, SCB Stockholm 1980

Zarkovich S S: Quality of statistical data, Food and Agriculture Organization of the United Nations, Rome 1966, kap 2.1

## 3.1.2 Konsistensstudier

Konsistensstudier syftar till att se efter hur uppgifter som samlats in i en undersökning förhåller sig till allmänt accepterade kunskaper om de variabler som ingår och deras förhållande till varandra.

Metoden utgår från ämneskunskaper från det område statistiken behandlar. Några exempel:

- Antal gifta män i befolkningen skall vara nästan exakt lika med antalet gifta kvinnor (i en monogam kultur)
- Bytesbalansens netto plus kapitalbalansens netto skall motsvara valutareservens förändring
- Antalet personer föremål för någon typ av vård i slutet av ett visst år plus nyintagna nästa år minus utskrivna nästa år skall motsvara antal personer föremål för vård i slutet av nästa år.

I den mån konsistensen avser uppgifter från samma undersökning (s k intern konsistens) kan dessa studier i stället utföras som kontroller i ett löpande granskningssystem.

Konsistensen behöver inte avse rent algebraiska identiteter utan kan också gälla sannolika (rimliga) relationer mellan värden på olika variabler. Vissa yrken är t ex regionalt koncentrerade. Om man i resultatredovisningen från en folkräkning fann ett stort antal järngruvearbetare i Malmö kommun vore det förmodligen idé att undersöka dessa närmare. Att hitta metoder för konsistensstudier kräver en del fantasi och uppfinningsrikedom. I de flesta undersökningar finns dock säkerligen sådana möjligheter, om man anstränger sig en smula.

### Referenser:

Morgenstern O: On the accuracy of economic observations, Second Edition, Princeton University Press, Princeton, New Jersey, 1965, s 132.

Zarkovich S S: Quality of statistical data, Food and Agriculture Organisation of the United Nations, Rome 1966, kap 2.2.

### 3.1.3 Preliminära - definitiva resultat

I de fall då preliminära resultat redovisas, baserade på t ex ofullständig granskning/rättning eller på ofullständigt uppgiftslämnande och definitiva resultat senare redovisas, kan man studera skillnaderna mellan de preliminära och de definitiva resultaten under en lång tidsperiod. På så sätt kan man få ett mått på en del av felen i de preliminära resultaten. En sådan studie kan göras för olika variabler och redovisningsgrupper.

Den skillnad som erhålls genom en jämförelse av detta slag blir ett mått på den sammanlagda effekten av de fel som finns i den preliminära statistiken men inte i den definitiva.

### Referens:

Byggnadsverksamheten i lantbruket 1977-1978, SM J 1979:12, sid 6.

### 3.2 Evalvering av fel på objektsnivå

Fel på objektsnivå genereras av felaktigheter vid insamlingen och bearbetningen av data för undersökningsobjekten. Felen kan vara av såväl slumpmässig som systematisk karaktär. Olika evalveringsmetoder kommer åt olika delar av felen. Med vissa metoder är det också möjligt att uppskatta de samlade effekterna av felen på objektsnivå. Nedan presenteras några olika metoder att evalvera felen på objektsnivå, varvid det inledningsvis anges vilka felkomponenter som metoden är inriktad på.

### 3.2.1 Jämförelser med annat material på objektsnivå

Metoden syftar till att mäta eller bedöma det sammanlagda (systematiska + slumpmässiga) felet på objektsnivå.

Jämförelser kan göras på total- eller urvalsbasis mellan två undersökningar/register, som innehåller samma objekt och har mist en gemensam undersökningsvariabel ('mäter samma sak'). Jämförelsen sker genom att värdena för samma objekt i de bägge materialen jämförs.

Om den undersökning eller det register man stämmer av mot kan antas ha avsevärt högre noggrannhet än den som är föremål för evalvering, kan den konstaterade (netto-) differensen tas som utgångspunkt för en skattning av det sammanlagda felet på objektsnivå (se vidare avsnitt 2.4 om felmått).

I övriga fall får man tillgripa rekonciliering (sammanjämkning) för de objekt där värdena i de två undersökningarna skiljer sig åt. Rekonciliering innebär att ett "sant" värde fastställs genom förnyad kontakt med uppgiftslämnarna, expertomdöme, omkodning eller studium av källmaterial, allt efter omständigheterna i den speciella undersökningen.

De mest omfattande evalveringsundersökningar av detta slag, som har gjorts vid SCB, gäller Folk- och bostadsräkningarna 1970 och 1975. Här jämfördes t ex uppgifterna som lämnats till FoB med dem som lämnats i arbetskraftsundersökningarna. Vid avvikelser mellan uppgifterna togs förnyade kontakter med uppgiftslämnarna.

#### Referenser:

Folk- och bostadsräkningen 1970. Resultat från evalveringsstudierna avseende sysselsättning och utbildning, SM Be 1974:3.

Folk- och bostadsräkningen 1975 del 9:2. Resultat från evalveringsstudie avseende sysselsättningsdata, SOS, Stockholm 1978.

Lundholm G: Registeruppgifter om socialhjälp - en jämförelse med annan statistik, SCB Rapport KEX 2011, 1976-02-18

### 3.2.2 Kontroll av uppgiftslämnande vid källan/återintervjuer med bättre mätmetod

Metoden är inriktad på det sammanlagda felet på objektsnivå. Huvudsakligen syftar den till att studera uppgiftslämnarfelet men genom att jämföra källmaterialet med det vid SCB bearbetade (t ex ADB-register) kan man dessutom kontrollera gransknings-, kodnings- och dataregistreringsmomenten.

En källkontroll innebär att en kvalitetskontrollant uppsöker uppgiftslämnare (normalt ett urval) i den aktuella undersökningen och ber att få ta del av det källmaterial som använts vid uppgiftslämnandet. Utifrån detta fastställer kvalitetskontrollanten vilken uppgift som borde ha lämnats och denna uppgift tas så som det sanna värdet för objektet. Man bortser härigenom dels från fel i källmaterialet dels eventuella fel vid kvalitetskontrollantens bearbetning. Tillvägagångssättet förutsätter ett skriftligt källmaterial och lämpar sig därför främst för undersökningar, där uppgiftslämnarna är myndigheter, företag, kommuner e d. Det skriftliga källmaterialet kan t ex vara register, akter eller bokföring.

När skriftligt källmaterial saknas (t ex i de flesta individundersökningar) får i stället återintervjuer med uppgiftslämnarna tillgripas. I de fall återintervjuerna genomförs med mätmetoder, som är bättre än dem i huvudundersökningen kan resultaten från återintervjuerna operationellt antas som sanna värden. Sådana mätmetoder kan t ex vara intervjuer i stället för postenkäter, expertintervjuare, mer omfattande frågebatteri, extra utbildning av intervjuare m m.

Kontrollerna kan ha flera syften. Det vanligaste är att skatta det sammanlagda felet på objektsnivå, vilket förutsätter ett sannolikhetsurval av viss storlek. Men om syftet begränsas till att få grepp om de problem uppgiftslämnarna har och förbättra frågeställningar, definitioner och anvisningar räcker det ofta med små subjektiva urval.

En variant av källkontroll kan användas när samma mätvärden finns tillgängliga på flera håll. Det kan t ex vara fråga om att undersökningsobjektet är skyldigt att lämna uppgifter till kommunala organ eller att ett värde finns tillgängligt både hos en leverantör och en köpare. Kontrollen består då i, att man samlar in uppgifter för en del av objekten hos den källa som inte svarar för uppgifterna till huvudundersökningen.

#### Referenser:

Bergman L R och Thorslund M: Svarkvalitén i AKU. Resultat från två studier med återintervjuer. SCB I/UI Metodproblem i individ- och hushållsstatistik nr 13.

Dalén J: Kvalitetskontroll av socialvårdsstatistik. Jämförelse mellan SCBs registeruppgifter och de sociala nämndernas beslutsprotokoll, Stencil SCB 1976-01-22

Friberg R: Industrins investeringar. Ett försök med återintervjuer för att mäta storleken av uppgiftslämnarfelet. SCB Rapport KEX 1010, 1973-08-15.

Friberg R: Omsättningen inom vissa tjänstenärings. Ett försök med återintervjuer för att mäta storleken av uppgiftslämnarfelet, SCB Rapport KEX 1012, 1973-08-14

Friberg R: Uppgiftslämnarfel i inkvarteringsstatistiken - kvalitetskontrollstudie KEX 2010, SCB Rapport, 1974-07-09

Rosenlund S: KEX 2012. Uppgiftslämnarfel i månadsstatistiken för industriarbetare, SCB Rapport 1974-07-03.

### 3.2.3 Indirekta kontroller av uppgiftslämnandet

Genomgång av källmaterial är ofta förhållandevis dyrbar. Man kan då i stället välja någon form av indirekt kontroll. Nedan nämns några - ganska oprövade- förslag till sådana kontroller som kan vara aktuella i vissa fall. Kontrollerna kan också tänkas fungera i ett granskningsammanhang.



Man kan låta uppgiftslämnarna (eller en del av dem) försöka ange något mått på uppgifternas osäkerhet och/eller en kortfattad beskrivning av hur den lämnade uppgiften har erhållits eller beräknats. I det senare fallet bör frågorna formuleras så, att svaren ger en uppfattning om hur väl de lämnade uppgifterna ansluter till SCBs definitioner. Detta kan ge en uppfattning om förekomsten av systematiska fel och kanske också ge uppslag till förändringar av frågeställningarna.

Om uppgiftslämnarna till följd av tidspress eller avsaknad av säkra uppgifter tvingas uppskatta vissa variabelvärden kan man be dem markera detta och vid nästa insamlingstillfälle begära in uppgifter om reviderade värden för den föregående undersökningen. Effekten på skattningarna kan på detta sätt beräknas.

Om man har starkt säsongberoende data kan man jämföra uppgifter på objektsnivå med motsvarande uppgifter på aggregerad nivå under en följd av statistikperioder. Om ett enskilt objekt uppvisar stora avvikelser i "mönstret" (låg korrelation med totalsumman) kan detta vara ett tecken på fel.

På ett liknande sätt kan man betrakta olika variabler, som normalt har en stark samvariation (hög korrelation). Variablerna kan härstamma från samma eller olika undersökningar. Man kan antingen utgå från korrelationen över tiden på aggregerad nivå eller från korrelationen mellan olika objekt vid ett och samma undersökningstillfälle. Om ett visst objekt faller ur "mönstret" kan detta tyda på ett fel.

Till sist skall även en metod som kallas frekvensanalys nämnas, som kan avslöja en viss typ av systematiskt uppgiftslämnarfel. Metoden innebär, att man undersöker, om vissa värden är speciellt vanligt förekommande i förhållande till vad man kan förvänta sig. Det kan t ex vara fråga om avrundningar till jämna 10-tal eller en anhopning av variabelvärden kring ett "idealresultat" eller ett genom lag föreskrivet gränsvärde (max eller min). Analysen kan enklast göras genom att man undersöker antalet noteringar för olika variabelvärden och

prickar in dessa i ett diagram. Felbilden kan fås fram under förutsättning av ett antagande om ungefär lika frekvenser för intilliggande värden, t ex genom att betrakta det sanna värdet som ett glidande medelvärde av frekvenserna för ett antal intilliggande värden.

#### Referens:

Zarkovich S S: Quality of statistical data, Food and Agriculture Organization of the United Nations, Rome 1966, kap 9.6.

#### 3.2.4 Replikering

Metoden leder till en objektiv uppskattning av det slumpmässiga felet i ett visst moment av en statistisk undersökning. Man kommer däremot inte åt den systematiska felkomponenten.

Metoden innebär att man för ett urval av objekten i undersökningen upprepar ett visst moment i produktionsprocessen. Det kan röra sig om återintervjuer av personer, vilket bl a leder till ett mått på svarsvariationen. Det kan också vara fråga om ett visst bearbetningsmoment såsom granskning, kodning eller dataregistrering, som upprepas, vilket leder till ett mått på variabiliteten i just detta moment.

De matematiska beräkningar som erfordras finns redogjorda för t ex i Cochran, 1977. Vi visar här bara det enklaste fallet med en variabel som endast antar två värden (här kallade 0 och 1). Nedanstående tablå visar antalet objekt med olika kombinationer av värden i dels den ursprungliga undersökningen och dels replikeringen.

Replikering

		0	1	Summa
Under- sökning	0	a	b	a + b
	1	c	d	c + d
Summa		a + c	b + d	n

Värdena i huvuddiagonalen (a och d) står sålunda för de fall, där uppgiftslämnarna lämnat samma uppgift vid bägge tillfällena.

Variabiliteten uttrycks genom  $\frac{b + c}{n}$ , medan nettoskillnaden blir  $\frac{b - c}{n}$ . Nettoskillnaden kan ses som ett mått på den osäkerhet i resultaten som variabiliteten leder till.

En strikt analys av resultaten från en replikeringsundersökning förutsätter att någon korrelation mellan felen för olika objekt inte föreligger och att mätningarna vid de två tillfällena har samma bakomliggande varians. Dessutom förutsätts oberoende mellan undersökningstillfällena. Vid t ex intervjuundersökningar kan en korrelation antas föreligga mellan mätfelet för samma intervjuares undersökningsobjekt. I sådana situationer krävs mer komplicerade undersökningsförfaranden inkluderande även interpenetrering. Interpenetrering innebär att man delar in ett urval i suburval på så sätt att ingen korrelation föreligger mellan objekt i olika suburval. Exempelvis kan ett suburval bestå av en intervjuares samtliga undersökningsobjekt.

Vid återintervjuer medför förutsättningen om oberoende att tidsskillnaden mellan undersökningen och återintervjun bör vara så stor som möjligt. Å andra sidan uppstår minnesfel vid en stor tidsförskjutning. Det är därför en intrikat uppgift att hitta den optimala balansen mellan dessa två motstridiga krav. Metoder för att göra mätvärdena oberoende trots liten tidsskillnad kan vara att variera ordningsföljden mellan frågorna eller deras formulering för att få intervjupersonen att tänka i nya banor (men så att han ändå besvarar samma sak).

#### Referenser:

Cochran W G (1977): Sampling techniques, Third edition, Wiley, New York, kap 13.

Dalenius (1974): Ends and means of total survey design, Forskningsprojektet FEL I UNDERSÖKNINGAR, Stockholms universitet, Statistiska institutionen, kap XIII

Wikman A (1980): Svarsprecisionen i surveyundersökningar om levnadsförhållanden, Metodproblem i individ- och hushållsstatistik nr 14, SCB I/UI.

### 3.2.5 Evalvering av kodning

Här ges en kort sammanfattning av de metoder, som utarbetats vid SCB för kontroll av noggrannheten i kodningsmomentet. En omfattande redogörelse finns i Harvig och Lyberg (1977).

Gemensamt för metoderna är någon form av upprepning av kodningsproceduren för en mängd blanketter, varvid de erhållna värdena jämförs. En strävan bör vara att uppnå en (åtminstone operationellt definierad) sann kod. Därigenom kan nämligen det sammanlagda felet i kodningsmomentet uppskattas och brutto- och nettofel beräknas liksom vid motsvarande evalvering av uppgiftslämnarfelet (avsnitt 3.2.2 ovan).

En sann kod kan genereras på flera olika sätt. Om man har tillgång till en kodare, vars skicklighet klart överstiger de normala kodarnas (en s k expertkodare) kan dennes kod definieras som sann. Om minst tre kontrollkodare kodar samma blankett kan man använda en majoritetsregel för att fastställa den sanna koden. Eljest kan någon form av rekoncilieringsförfarande tillgripas när den ursprungliga kodarens (primärkodarens) kod inte överensstämmer med kontrollkodarens.

Om möjligheter saknas att fastställa en sann kod får man i stället studera olika typer av kodarvariabilitet via ett replikeringsförfarande motsvarande det i 3.2.4 ovan. Man skiljer därvid mellankodarvariabilitet från inomkodarvariabilitet. Det förra är ett mått på variationen mellan olika kodares kodning av samma blankett medan det senare mäter samma kodares olika kodning vid olika tillfällen.

#### Referenser:

50 brottskoder och deras innehåll, Kvalitetskontroll av polisstatistiken, Promemorior från SCB 1975:13

Harvig H och Lyberg L (1977): Kontroll av kodning, Statistiska centralbyrån, intern handbok.

Harvig H: Kontrollkodningsexperiment på blanketter för in-skrivningsuppgifter till högre studier, SCB Rapport KEX 1003, 1973-10-22

Harvig H: Kontrollkodning av dödsbevis, SCB Rapport KEX 1004, 1973-11-20

Lyberg L, Nordling P och Elmdahl J: Kodningskvaliteten i läraryregistret - KEX 1002. SCB Rapport 1973-08-21

Lyberg L: Kodningskvaliteten i FoB 70 - ekonomisk aktivitet. Lägesrapport VI, EU-VI, Stencil SCB, 1974-02-04

Lyberg L: Kodningsmodeller för oberoende kontrollkodning. SCB Rapport KEX 1015, 1974-01-29

### 3.2.6            Evalvering av stansning

Stansning (dataregistrering) är normalt en mycket liten felkälla i jämförelse med andra felkällor i statistikproduktionen. Därför är arbetet som lagts ner på att evalvera denna felkälla inte heller stort. (Den löpande kontrollen är däremot av tradition relativt omfattande.)

I ett fall har emellertid omfattningen av felen och dess effekter på de statistiska resultaten studerats. Metoden var här att jämföra primärstansade poster med kontrollstansade via maskinell matchning och även ta fram resultattabeller baserade på båda materialen och jämföra dessa.

#### Referenser:

Haglund T och Holgersson M: Dödsorsaksstatistiken vid I/RS; evalveringsstudie av stansningen och den maskinella granskningen. Redovisning av resultat. SCB P/STM, DELRAPPORT RIS-I/RS 1, 1980-01-10

Prov rörande stansfelens omfattning och betydelse för en statistikprodukt (UVAV). Kvalitetskontrollmetoder vid stansning, slutrapport från projektet KOST, Bilaga 2, 1975-07-11

### 3.2.7 Evalvering av granskning/rättning

Granskning och rättning av inkomna statistikuppgifter har i förhållande till andra produktionsmoment den säregenheten att det syftar till att eliminera fel från andra produktionsmoment. För detta syfte byggs system upp med kontroller utifrån definitioner på och relationer mellan olika variabler.

Det är inte alls självklart vad en evalvering av granskning/rättning skall syfta till. Flera mål är tänkbara:

- Att mäta effekterna av granskning/rättning på de redovisade resultaten. Detta uppnås t ex genom att jämföra tabeller baserade på ogranskat material med dem som baserats på granskat material.
- Att utreda hur stor del av den totala felmängden som man lyckats rätta till i gransknings-/rättningsystemet. Detta förutsätter att man lyckats uppskatta den totala felmängden först.
- Att kontrollera de fel som granskningspersonalen gör i förhållande till givna instruktioner
- Att mäta nettoeffekten på noggrannheten av granskning/rättningsarbetet
- Att analysera de fel, som upptäcks i granskningen som ett led i uppskattningen av andra felkällors betydelse

Granskning och rättning är en stor kostnadspost i de flesta SCB-undersökningar. Det är därför anmärkningsvärt att så litet resurser lagts ner på att evalvera effekterna av detta arbete.

Referens:

Eiderbrant-Nilsson G: Försök med oberoende granskning av  
inkommande uppgifter om kommunernas finanser - KEX 1013.  
SCB Rapport, 1974-10-21

### 3.3            Evalvering av fel på populationsnivå

I detta avsnitt skall främst det systematiska felet på populationsnivå behandlas. Det slumpmässiga felet är det som uppstår till följd av urval plus de mindre tillskott till detta som kommer från det faktum att vid undertäckning och bortfall ett färre antal objekt med mätvärden återstår. Detta kan beräknas inom ramen för den traditionella samplingteorin.

Vi skall alltså här i huvudsak begränsa oss till bortfallets och täckningsfelets snedvridande effekt. (Det finns ytterligare en typ av systematiskt fel som härrör från uppräkningsformeln som sådan men detta finns behandlat i läroböckerna i förekommande fall.) Situationen i icke-surveystatistik skall också uppmärksammas. Avslutningsvis skall vi ta upp några typer av bearbetningsfel, som uppträder på populationsnivån.

#### 3.3.1            Jämförelser med sant värde

Metoden leder till en uppskattning av det systematiska felet som härrör från bortfall och/eller undertäckning i surveys.

Metoden utgår från att det är möjligt att få tillgång till ett sant värde för de objekt som saknas. Detta kan tänkas i följande situationer:

1) I de fall man väljer att publicera preliminär statistik baserat på ett avsevärt större bortfall än den definitiva statistiken. Genom att jämföra de två statistikversionerna får man då ett mått på de effekter som den del av bortfallet hade, som så småningom kom in med uppgifter. Det bortfall som kvarstår även i den definitiva statistiken kommer man dock inte åt på detta sätt. Dessutom bör det observeras att det mått som tas fram på detta sätt också är utsatt för slumpmässiga variationer.

2) Om undertäckningen i en undersökning beror på bristande aktualitet i ramen kan man vid en senare tidpunkt erhålla uppgifter från undertäckningsobjekten. I den mån man inte har tid att inkludera dessa i resultatredovisningen kan de i stället användas i evalveringssyfte.

3) Värdena för de saknade objekten eller eventuellt en del av dessa kan finnas att tillgå i en annan undersökning, ett register eller på annat sätt. I sådana situationer skiljer sig ofta definitioner åt en smula, uppgifterna kanske inte avser exakt samma tidpunkt. Dessa faktorer måste tas med i beräkningen vid denna typ av jämförelse.

#### Referens:

Kristiansson K-E: Bortfallsstudie i AKU. Metodproblem i individ- och hushållsstatistik nr 10, SCB I/UI, 1980-01-29

#### 3.3.2 Strukturjämförelser mellan svarande och icke-svarande

Metoden är närmast tillämplig på bortfall i surveys men kan också tänkas användas i undertäckningssituationer.

Idén är att försöka ta reda på om svarsgruppen och bortfallsgruppen (ev undertäckningsgruppen) skiljer sig åt beträffande sådana bakgrundsvariabler som är korrelerade med undersökningsvariablerna, t ex ålder, kön eller bostadsort. Om detta är fallet kan en snedvridande effekt till följd av detta antas. Från handboken Räkna med bortfall (SCB 1980) hämtar vi följande exempel från en amerikansk undersökning - Fink och Crossby (1951) - med 31 % bortfall.



	Svars- grupp (%)	Bort- falls- grupp (%)
Andel män	46	46
Sysselsättning		
Intellektuella yrken	7	7
Företagare m fl	8	6
Kontorsanställda	14	12
Arbetare	15	15
Serviceyrken	5	7
Hemmafruar	31	22
Övriga	20	31
Telefoninnehav	87	81
Villa	51	34

I detta fall finns för vissa variabler stora skillnader mellan svarsgrupp och bortfallsgrupp. I den mån dessa variabler är korrelerade med mätvariablerna resulterar detta i en snedvridande effekt på undersökningsresultaten. Det bör observeras att strukturjämförelser aldrig kan fria, bara fälla. Även om dessa uppvisar små skillnader kan en snedvridande effekt mycket väl förekomma.

#### Referenser:

Fink R och Crossby H M: Response and nonresponse in a probability sample, Int Journ Opin Att Res, 1-19

#### 3.3.3 Studier av variationer över svarsomgångar

Metoden är enbart tillämplig på bortfall i surveys. Datainsamlingen bör ha skett genom postenkäter.

Metoden går ut på att dela in de erhållna uppgifterna i grupper uppdelade efter när uppgifterna kommit in: omedelbart efter enkätens utsändning efter en påminnelse, efter två påminnelser etc. Man kan också dela in tiden efter blankettutsändningen i perioder. Vi visar här ett exempel på en sådan situation. Exemplet - från Finkner (1950) - gäller persikoodlare i North Carolina 1946:

	Antal	% av	genomsnittligt antal
	odlare	population	fruktträd per odlare
Svarande utan påminnelse	300	10	456
Svarande efter påminnelse	543	17	382
Svarande efter två påminnelser	434	14	340
Bortfall efter två påminnelser	1 839	59	290
Hela populationen	3 116	100	329

Situationen var här den att för en undersökningsvariabel - antal fruktträd - fanns en uppgift tillgänglig på populationsnivå. Därigenom kunde en trend konstateras, såtillvida att odlare med få fruktträd svarade senare eller inte alls. Om man kan överföra denna trend på andra undersökningsvariabler (t ex persikoskörd) kan man uppskatta bortfallets effekt på skattningarna av dessa variabler (eller göra en bortfalls-korrigerings).

Normalt finns dock inte tillgång till någon information om hela populationen. För att en sådan här uppställning då skall vara till någon hjälp måste man ur denna kunna dra någon slutsats om det som kommer efter sista svarsomgången - bortfallet. Men någon objektiv grund för någon "extrapolering av trenden" finns inte. Det blir i så fall en subjektiv bedömning av statistikproducenten.

En bättre grund för en sådan extrapolering har man om man också genomför strukturjämförelser mellan svarsomgångarna och kan konstatera en motsvarande trend för sådana bakgrundsvariabler, som av saklogiska skäl kan antas vara korrelerade med undersökningsvariabeln.

#### Referenser:

Blunt A, Claesson L och Lindström H: Bortfallsfel - en sammanställning av fallstudier. SCB P/STM Stencil, 1977-08-16

Finkler A L (1950): Methods of sampling for estimating commercial peach production in North Carolina. North Carolina Agricultural Experiment Station Technical Bulletin, nr 91.

### 3.3.4 Jämförelser mellan skattningsmetoder

Metoden syftar till att bedöma bortfallsfelet när bortfalls-kompenserande åtgärder vidtagits.

I många undersökningar med bortfall försöker man att kompensera för det fel som detta orsakar genom olika förfaranden såsom kompensationsvägning och imputering. En redogörelse för dessa förfaranden finns t ex i handboken Räkna med bortfall (SCB 1980). Man har då möjlighet att jämföra resultaten av dessa olika skattningsmetoder med varandra och bedöma dem sinsemellan. Vi visar här ett exempel på detta där även ett sant värde kunnat tas fram. Exemplet är hämtat från Sollander m fl (1974), som också innehåller en närmare redogörelse för innebörden av kompensationsmetoderna.

Tabell Röstningspartisympatifördelning, efterstratifiering, lika medelröstningssannolikhet. Procent

	Val- resul- tat	Skattningar sedan bortfallet fördelats		
		i) som svars- individerna	ii) med imputa- tioner	iii) på parti och strata
C	25,1	23,6	23,3	22,6
F	9,4	11,0	11,6	12,7
M	14,3	11,8	11,3	11,0
S	43,6	45,6	45,8	45,5
VpK	5,3	5,1	5,0	5,2
Småpartier	2,3	3,0	3,1	3,0
Summa	100	100	100	100
Summa absoluta avvikelser		8,5	10,3	11,8

När man har tillgång till det sanna värdet som här blir detta till synes en idealisk evalvering. Den kompensationsmetod som har den lägsta summan av absoluta avvikelser har det minsta totalfelet och är följaktligen den bästa. Haken är emellertid att om man bara har tillgång till det sanna värdet vid ett enda tillfälle kan man inte veta hur pass stabila avvikelserna är.

Normalt har man inte tillgång till något "sant" värde. Några bestämda slutsatser kan då inte dras från en sådan här jämförelse. Den kan dock ändå vara värdefull som utgångspunkt för en vidare analys av olika bortfallskompenserande förfarandens egenskaper.

#### Referens:

Sollander S, Wahlström S, Wärneryd B och Langlet P: Teori och metodik vid partisymptatiundersökningar, Forskningsrapport nr 5. SCB, avdelningen för uppdragsverksamhet, 1974.

#### 3.3.5 Matchning av register

Metoden syftar till att avslöja täckningsfel i ramen för en survey eller i ett register och i bästa fall mäta under- och/eller övertäckningsfelets storlek.

Metoden går helt enkelt ut på att matcha den ram/det register (A) som används vid undersökningen med ett kontrollregister (B) som omfattar helt eller delvis samma population som A. I de fall inte båda registren är ADB-lagrade kan en matchning på urvalsbasis bli aktuell. De slutsatser som kan dras ur dessa matchningar är beroende på den tilltro man sätter till respektive register.

Om register B betraktas som ett sant register utgör de objekt som finns i A men inte i B (eller eventuellt en relevant delmängd av B) övertäckning i A. Objekt som finns i B (ev delmängd) men inte i A utgör undertäckning i A.

Om inget av registren kan betraktas som sant i förhållande till det andra måste de konstaterade avvikelserna göras till föremål för ytterligare granskning - rekonciliering. Liksom i 3.2.1 ovan kan denna vara av olika slag; central expertbedömning eller inhämtande av ytterligare uppgifter om objektet kan bli aktuellt.

#### Referenser:

Dalén J: Kvalitetskontroll av socialvårdsstatistik. Sambearbetning mellan fosterbarnsregistret och registret över barn med beslut om åtgärder enligt barnavårdslagen. SCB, I/RS, Stencil, 1976-05-05.

Dalén J: Evalvering av tjänsteenkäten. Plan för undersökning av eventuella populationsskillnader mellan CFR och riksbankens valutaanmälningar. SCB, P/STM, Stencil, 1980-06-17.

#### 3.3.6 Analys av eftersläpningseffekter i register

Metoden är inriktad på att avslöja fel som beror på ett registers uppdateringssystem (se avsnitt 1.6 ovan).

Ett register uppdateras alltid med viss fördröjning. Ett objekts födelse, död eller förändring registreras inte omedelbart utan efter en viss tid, varierande från register till register och från fall till fall. Därigenom genereras fel, som behåftar registret vid varje bestämd tidpunkt.

Under förutsättning att den verkliga tidpunkten för uppdateringens ikraftträdande är känd kan nu eftersläpningseffekterna studeras. Detta kan ske på olika sätt. Man kan t ex tänka sig att beräkna eftersläpningstidens fördelning och därifrån skatta de fel detta medför i registret. Eller man kan matcha två i tiden näraliggande registerversioner för att se vilka av uppgifterna i den senare versionen som borde ha ingått även i den tidigare. Det senare förutsätter att registeruppgifterna är daterade. Även andra tillvägagångssätt är tänkbara.

På det här sättet kan felen i registret mätas i en följd av registerversjoner. Man kan då eventuellt konstatera en viss stabilitet i felnivån, som kan antas fortsätta in i framtiden.

En typ av eftersläpning är svårare att kontrollera på detta sätt nämligen rapporteringen av ett objekts död. När den har registrerats i kommande registerversjon ingår ju inte objektet längre. Här får man i stället tänka sig en annan metod, t ex matcha en fil med "döda" objekt mot den registerversjonen vi vill kontrollera. I vissa fall kan man också tänka sig att på urvalsbasis kontrollera kvarstående objekt från en registerversjon till nästa.

#### Referens:

Dalén J: Kvalitetskontroll av socialvårdsstatistik. Undersökningar av beslut om åtgärder enligt barnavårdslagen. SCB, I/RS, Stencil, 1975-07-23.

#### 3.3.7 Kontroll av ofullständig rapportering i händelsebaserad statistik

Som ovan (i avsnitt 1.3) nämnts bygger den händelsebaserade statistiken på att uppgiftslämnarna på eget initiativ lämnar en statistikrapport, när en viss händelse äger rum. Det är uppenbart att en sådan statistik är mycket känslig för fel beroende på glömska eller bristande rutiner hos den enskilde uppgiftslämnaren. Ofta är dock statistikrapporteringen i dessa fall samordnad med ett administrativt anmälningsförfarande, vilket är en hygglig garanti för fullständighet i rapporteringen.

Att kontrollera fullständigheten i dessa fall kan vara ganska besvärligt. En icke-bokförd händelse syns inte någonstans. Ett sätt att hålla kontroll på processen är att kontrollera varje uppgiftslämnarens rapportering i tiden. Oförklarliga fluktuationer kan då leda till närmare undersökning. Misstänker man att blanketter "kommer bort på vägen" kan en kontroll mot ett motsvarande administrativt material vara aktuell.

Den speciella problematik som mörkertalen utgör i brottsstatistiken (och även i vägtrafikolycksstatistiken) måste tacklas genom särskilda statistiska undersökningar av surveytyp, riktade till individer eller institutioner med kunskaper om det aktuella området. En sådan undersökning nämns i referensen nedan.

Referens:

Persson L G: Offer för tillgrepp, skadegörelse och våld - en redovisning av 1974 års offerundersökning. Promemorior från SCB 1977:7

### 3.3.8 Kontroll av bearbetningsfel på populationsnivå

Bearbetningsfel förekommer inte bara på objektsnivå. Man kan tänka sig flera olika typer av bearbetningsfel som ger sig till känna på populationsnivå. Här skall nämnas några typer av sådana fel och hur de kan upptäckas. Det är en praktisk fråga huruvida de "metoder" som nämns ingår i ett löpande granskningsarbete eller i en evalvering i efterhand.

Programmeringsfel är inte ovanliga i statistikproduktionen, men ofta svåra att upptäcka. Genom uttestning mot testmaterial (inklusive produktionsmaterial) kan stora fel normalt förebyggas. I efterhand kan t ex tabeller från olika tabellprogram konsistensgranskas och man kan jämföra med föregående statistikperiodstabeller. En mer omfattande diskussion om fel i data-system och -program återfinns i Statskontoret-SCB (1977).

Utskriftsfel upptäcks i allmänhet genom motläsning. Någon särskild efterhandskontroll är sällan aktuell.

Fel på grund av borttappade blanketter i gransknings- eller dataregistreringsprocessen kontrolleras t ex genom att blanketter in i varje moment räknas och jämförs.

Referens:

Statskontoret - Statistiska centralbyrån: Kvalitetsskydd av data, Liber Tryck, Stockholm 1977.