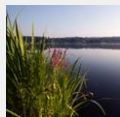


Bortfall – Konsekvenser Varför det kan vara allvarligt med bortfall.

Ann-Marie Flygare
Metodstatistiker, SCB





Konsekvenser av Bortfall

- Introduktion
- Illustration av hur bortfall påverkar resultaten i en statistisk undersökning – fiktiv undersökning
- Redovisning av bortfall



Konsekvenser av Bortfall

- Hur tänker du om bortfall vid statistiska undersökningar?
- Hur definierar du bortfall?
- Hur bedömer du resultat från statistiska undersökningar som har bortfall?
- När tycker du bortfall blir ett bekymmer?



Konsekvenser av Bortfall

- Ökad kunskap om konsekvenser av bortfall – hur det påverkar vår statistik.
- Är statistiken trovärdig?
- ”Statistikproducenten skall informera användarna om faktorer av betydelse för en korrekt tolkning av statistiken”

Policy för användarorienterad kvalitetsredovisning av statistik, MIS 1983:1 SCB



Konsekvenser av Bortfall

- **Definition:**

De element (objekt) i ramen som **tillhör** målpopulationen och man har planerat att undersöka, men som man ej fick något svar ifrån.

(Karin Dahmström, Från datainsamling till rapport – att göra en statistisk undersökning)

- Objektsbortfall
- Partiellt bortfall



Konsekvenser av Bortfall

- Färre svarande = Större slumpfel (felmarginal)... och det mest allvarliga med bortfall:
- Om de i bortfallet avviker från de svarande får vi ett snedvridet resultat → systematiskt fel
- Högre bortfall påverkar undersökningarnas kvalitet

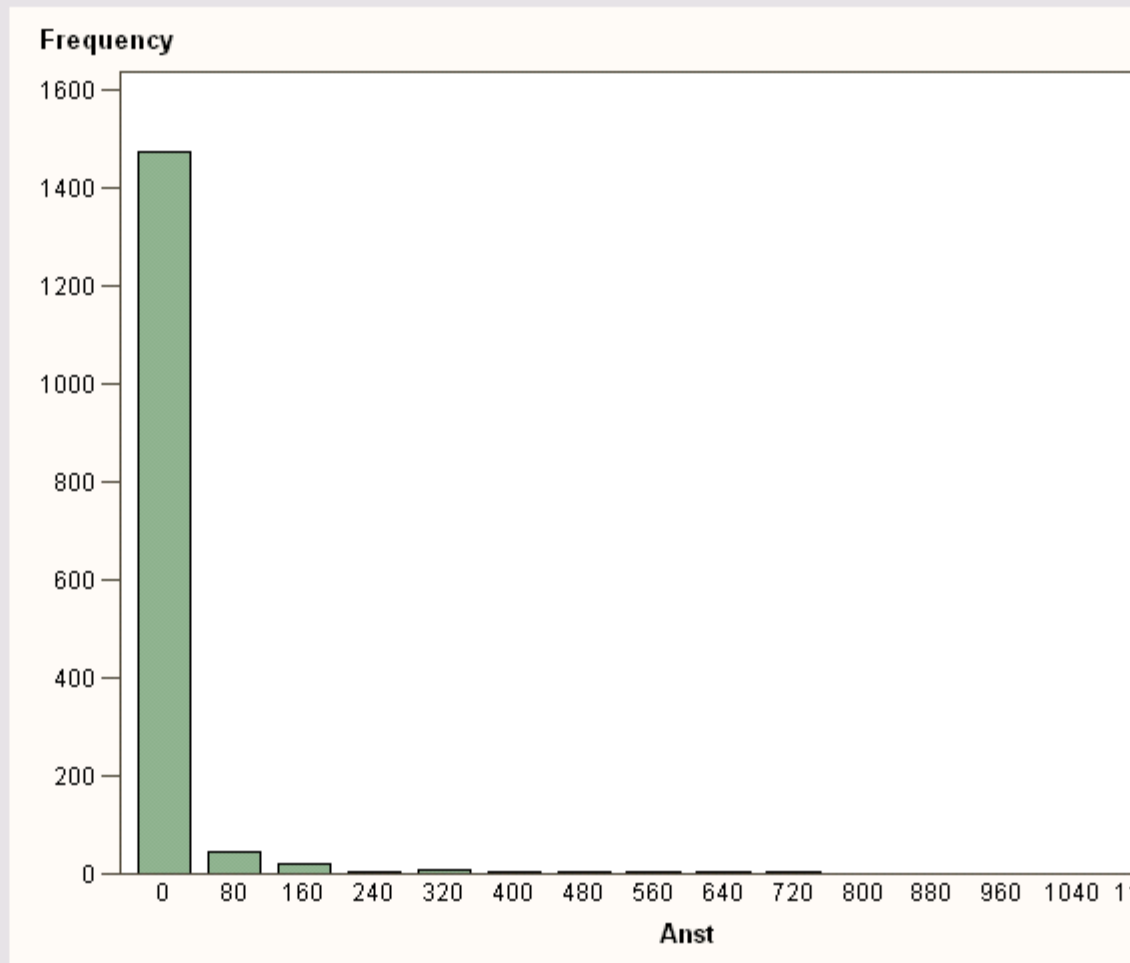


Fiktiv undersökning

- Population: Företag
N=1557 (antalet företag i populationen)
- Intresserade av parametern:
 - Genomsnittliga antalet anställda: $\mu=12.9$
- Skatta μ genom att dra ett OSU av företag (n=200)
- Resultat vid:
 - Inget bortfall
 - Slumpmässigt bortfall, olika andel bortfall
 - Systematiskt bortfall



Stapeldiagram



Genererad av SAS System ('Local', XP_PRO) den 01 mars

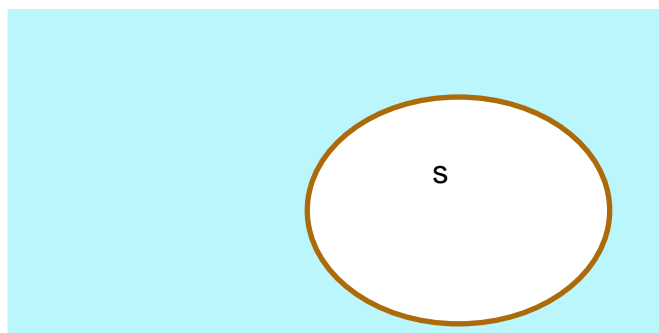
Fiktiv undersökning

- Genomsnittliga antalet anställda $\mu = 12.9$
- OSU av 200 företag.

Inget bortfall:

- Skattade genomsnittliga antalet anställda, $\bar{x}_s = 12.3$
intervallskattning (95%) : (6, 18.6).

U: alla objekt i populationen



s: objekt i urvalet



Fiktiv undersökning

- Resultat vid **slumpmässigt bortfall**, (10, 25, 50)%



Bortfall (%)	Antal svarande	Skattning	intervallskattning	Felmarginal
0	200	12.3	(6, 18.6)	6.3
10	180	12.5	(4.4, 20.5)	8.0
25	150	13.1	(4.2, 22.0)	8.9
50	100	13.0	(1.9, 24.1)	11.1

Fiktiv undersökning

- Systematiskt bortfall, inget företag av de 200 utvalda med fler än 50 anställda svarar på undersökningen:

Bortfall (%)	Antal svarande	Skattning	intervallskattning	Felmarginal
5	191	4.06	(2.2, 5.9)	1.84

- Slutsatser?



Fiktiv undersökning

	Population	Inget bortfall	Systematiskt bortfall	
			svarande	bortfall
Medelvärde	12.9	12.3	4.1	186.8
Standard avvikelse	58.6	48.7	13.8	136.7

313
434
87
75
50
125
189
320
88



Teoretiskt resonemang

- \bar{x} : Medelvärde, inget bortfall
- \bar{x}_s Medelvärde för svarande
- \bar{x}_b Medelvärde för bortfall
- w_s Andel svarande
- w_b Andel bortfall

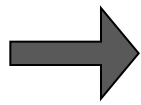
Avvikelsen mellan det observerade värdet \bar{x}_s och det okända värdet \bar{x} är ett mått på det systematiska felet som bortfallet orsakar:

$$\bar{x}_s - \bar{x} = \dots = w_b (\bar{x}_s - \bar{x}_b)$$



Konsekvenser av Bortfall

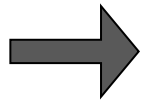
Varje objekt i urvalet har en känd urvalssannolikhet



Grunden för statistisk inferens är uppfylld

Bortfall förstör detta!

Sannolikheten att ett objekt ingår i den slutliga svarsmängden är produkten av urvalssannolikheten och sannolikheten att svara (som är okänd).



Stickprovet är troligen inte längre representativt

Redovisning av Bortfall

- Bristfällig redovisning av bortfall när resultat från statistiska undersökningar redovisas förekommer.
- Statistikrapporter bör innehålla redovisning av bortfall.
- Enhetlig redovisning





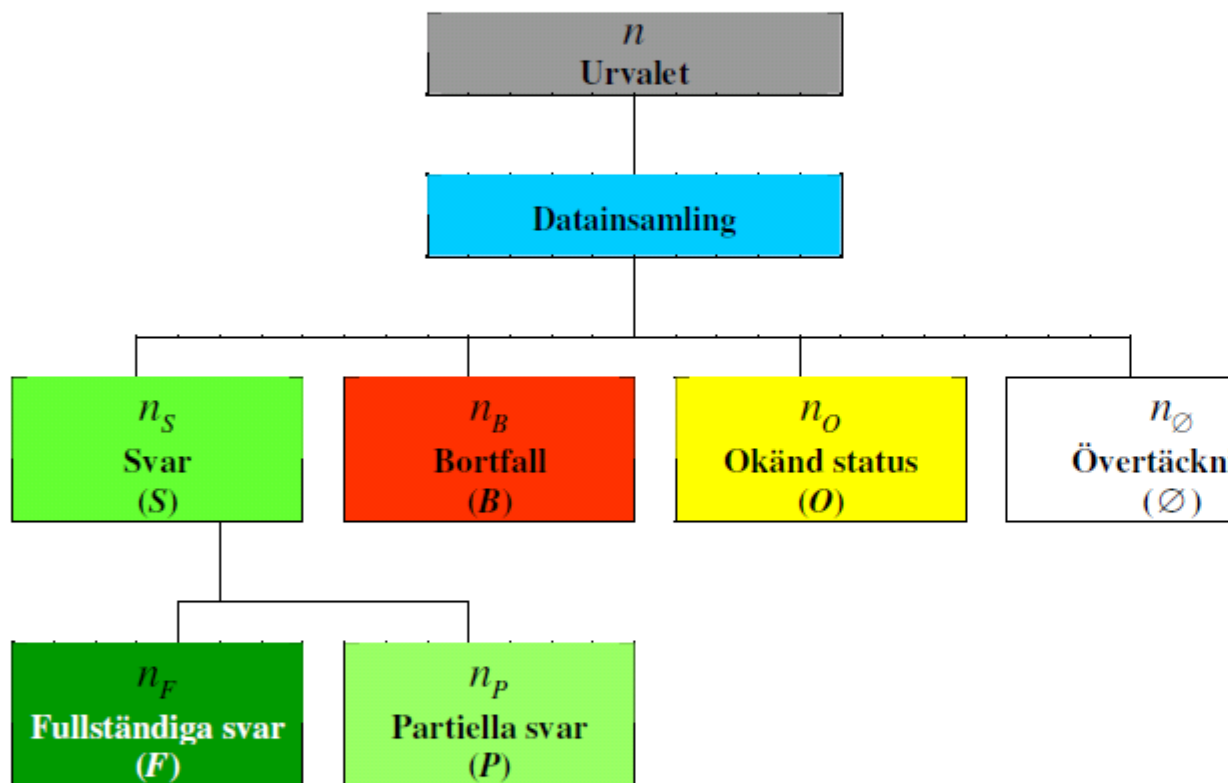
Standard för bortfallsberäkning

- Sektionen för Surveystatistik 2005
- Representanter från: Universitet, SCB, SMIF, FSM, ESOMAR

http://surveytest.mezzotech.se/?page_id=35



STANDARD FÖR BORTFALLSBERÄKNING



Redovisning av bortfall

- Ovägt bortfallsmått (objektsbortfall)

$$BA = 1 - \frac{n_s}{n_s + n_b + un_o}$$

- Vägt bortfallsmått:

$$BA = 1 - \frac{\sum_s d_k}{\sum_s d_k + \sum_b d_k + u \sum_o d_k}$$



Redovisning av bortfall

- Förslag till innehåll i en bortfallsredovisning:
 - Allmän redovisning:
 - Beskrivning av målpopulation i tid och rum, inklusive eventuella begränsningar (cut off)
 - Urvalsram
 - Insamlingsmetod, insamlingstidpunkt
 - Typ av urvalsförfarande
 - Uppgift om totala antalet element i urvalet



Redovisning av bortfall

- Redovisning av svar och bortfall:
 - Antal svarande med användbara observationer
 - Antal utvalda element med icke användbara observationer
 - Val av ett eller flera bortfallsmått och dess numeriska värden.
- Försök ange storleken på de systematiska fel som bortfallet kan ha fört med sig, samt beskriv de åtgärder som har vidtagits för att minska effekterna av dessa fel.



Redovisning av bortfall

- Jaga svarsprocent, kan vara mindre bra
- Bra eller dåligt:
”Vi redovisar inga resultat när svarsandelen är mindre än 60%”

Bortfallsanalys!



Slutsatser

- Bortfallsfel förekommer!
- Svarsprocenten i sig är ingen bra indikator på bortfallsfelet
- Utan studier av bortfallsfel är det svårt att veta ifall bortfallet är ett problem i en undersökning

Konsekvenser av Bortfall

- Hur tänker du om bortfall vid statistiska undersökningar?
- Hur definierar du bortfall?
- Hur bedömer du resultat från statistiska undersökningar som har bortfall?
- När tycker du bortfall blir ett bekymmer?

