



**Karolinska
Institutet**

Borde vi mäta statistisk osäkerhet vid totalundersökningar?

Paul Dickman

Sandra Eloranta

Therese Andersson

Institutionen för Medicinsk Epidemiologi och Biostatistik (MEB)

Karolinska Institutet, Stockholm

Det korta svaret är:

- JA! – om syftet med den statistiska undersökningen är vetenskapligt (analytiskt).
- Om syftet istället är administrativt finns inga skäl till att redovisa standardavvikelser, konfidensintervall etc.
- Oklarheten består inte i om osäkerhet ska redovisas utan hur den i själva verket ska mätas – dvs hur bör man härleda variansformler då målpopulationen är en högst hypotetisk superpopulation?

Överblick av dagens seminarium

- Presentation + kort introduktion till vår forskning.
- Vad vi, i teorin, har fått lära oss om slumpmässig variation i statistiska undersökningar
- Hur vi, i praktiken ser på slumpmässig variation inom registerbaserad forskning.
- Borde vi mäta statistisk osäkerhet vid totalundersökningar?
- Diskussion och frågor

Epidemiologiska studier för cancerpatientöverlevnad

- Epidemiologiska studier kan användas som ett verktyg för att utvärdera:
 1. effektiviteten i cancervården
 2. de åtgärder som satts in i syfte att minska konsekvenserna / påverkan av cancer
 3. identifiera potentiella riskfaktorer som kan kopplas till ökad dödlighet
- Studier av cancerpatientöverlevnad har ofta ett folkhälsoperspektiv och är av analytisk karaktär.

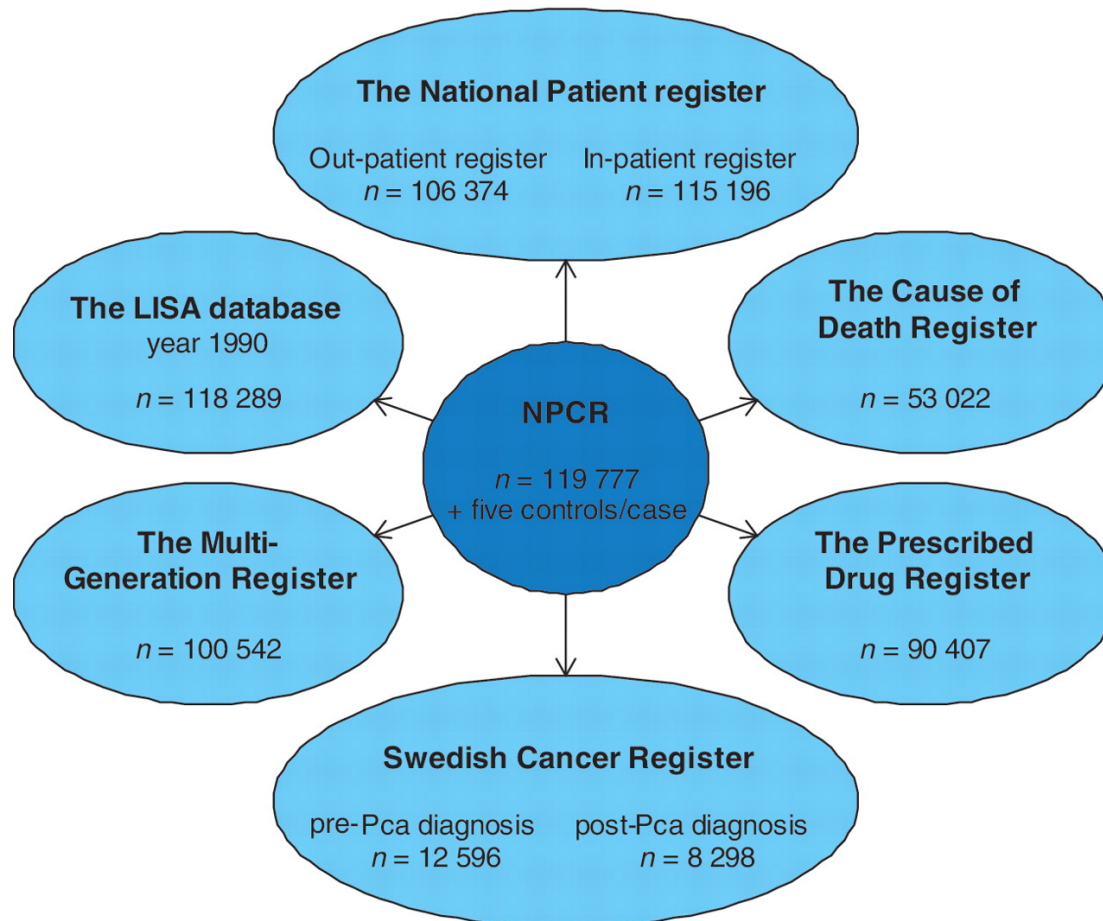
Populationsbaserade studier av cancerpatientöverlevnad

- Data från nationella/regionala cancerregister.
- Rutinberäkningar (exempelvis tidstrender i femårsöverlevnaden)
- Utveckling av existerande statistiska metoder för att skatta överlevnaden
- Särskilt intresse att utveckla kommunikationen av canceröverlevnadsstatistik.

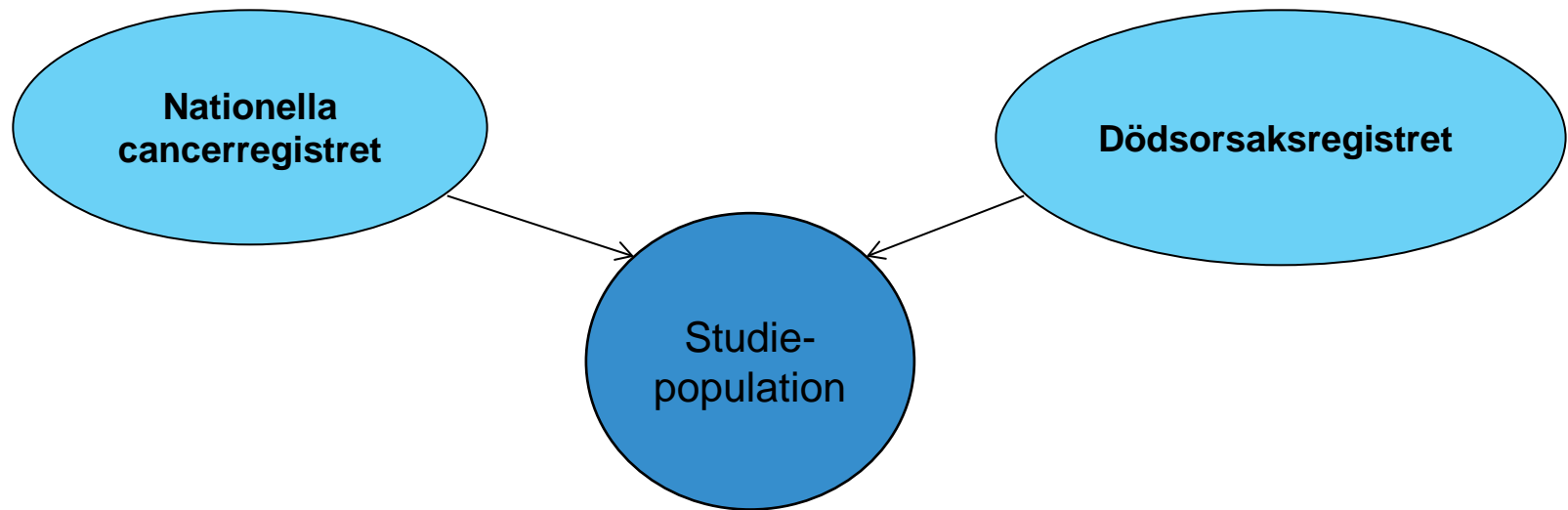
Ofta komplicerade databaser

(Van Hemelrijck et al. *Int. J. Epidemiol.* 2012)

PCBaSe Sweden 2.0



En enklare, men vanligare generell datastruktur





Success Story of Targeted Therapy in Chronic Myeloid Leukemia: A Population-Based Study of Patients Diagnosed in Sweden From 1973 to 2008

Magnus Björkholm, Lotta Ohm, Sandra Eloranta, Åsa Derolf, Malin Hultcrantz, Jan Sjöberg, Therese Andersson, Martin Höglund, Johan Richter, Ola Landgren, Sigurdur Y. Kristinsson, and Paul W. Dickman

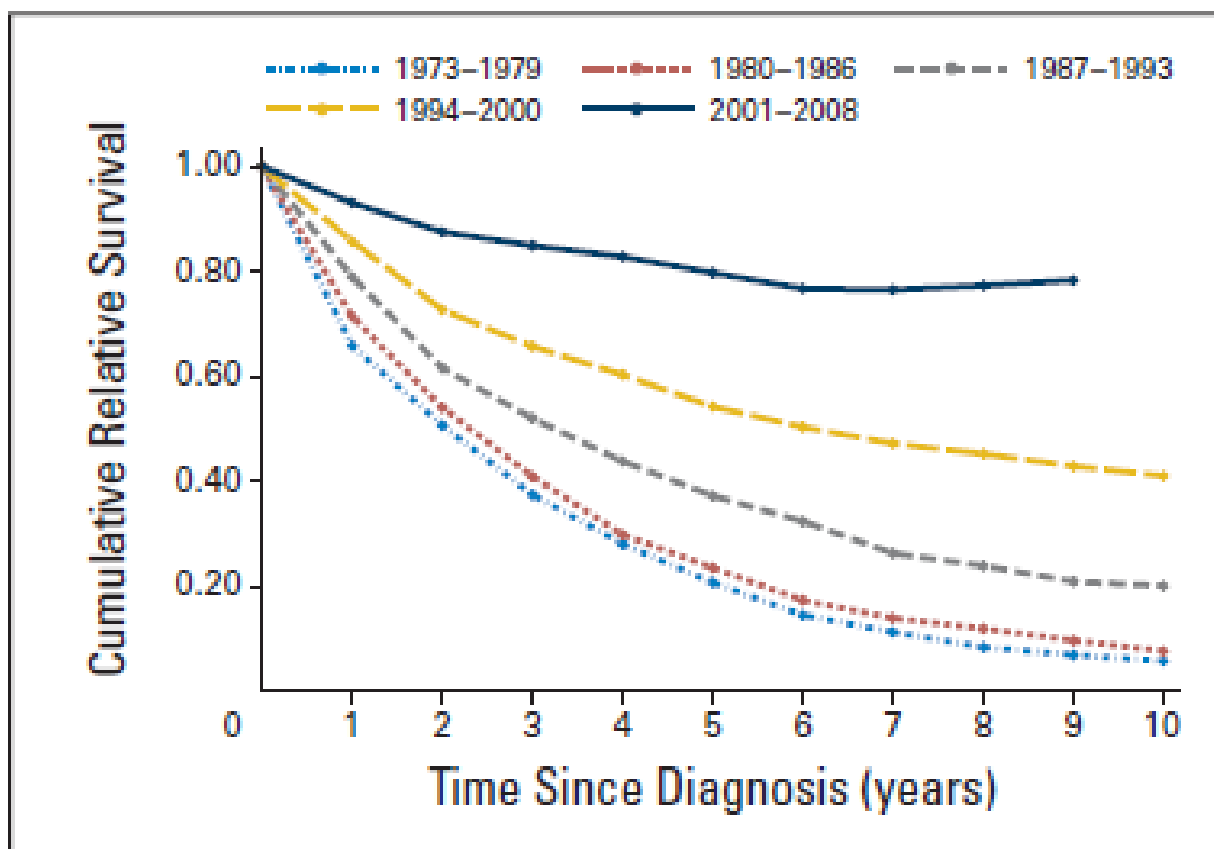


Fig 1. Relative survival ratios by calendar period of diagnosis.

1-årsöverlevnad per kalenderperiod och åldersgrupp

Table 3. Estimates of 1-, 5-, and 10-Year RS Among Patients With CML in Sweden Stratified by Age and Calendar Period of Diagnosis

Age (years)	1 Year									
	1973-1979		1980-1986		1987-1993		1994-2000		2001-2008	
	RS	95% CI	RS	95% CI	RS	95% CI	RS	95% CI	RS	95% CI
< 50	0.83	0.76 to 0.88	0.83	0.77 to 0.88	0.90	0.84 to 0.94	0.95	0.91 to 0.97	0.96	0.92 to 0.98
50-59	0.78	0.68 to 0.85	0.82	0.73 to 0.88	0.89	0.80 to 0.94	0.95	0.88 to 0.98	0.99	0.95 to 1.00
60-69	0.65	0.56 to 0.72	0.78	0.70 to 0.84	0.85	0.77 to 0.90	0.87	0.78 to 0.93	0.96	0.91 to 0.99
70-79	0.49	0.40 to 0.58	0.60	0.52 to 0.67	0.67	0.58 to 0.75	0.81	0.72 to 0.87	0.92	0.84 to 0.96
> 79	0.31	0.19 to 0.44	0.43	0.32 to 0.54	0.47	0.34 to 0.60	0.45	0.31 to 0.58	0.69	0.56 to 0.80

Läkartidningen 2004

Cancerforskare, sluta redovisa konfidensintervall när det inte behövs!

II Med förvåning kan man i vetenskapliga tidskrifter läsa svenska artiklar som redovisar osäkerhet i form av konfidensintervall eller p-värden trots att det av metodbeskrivningen framgår att undersökningen är en populationsbaserad totalundersökning som använt det svenska cancerregistret, ofta i kombination med andra heltäckande register som t ex registren över slutenvård eller folk- och bostadsräkningen.

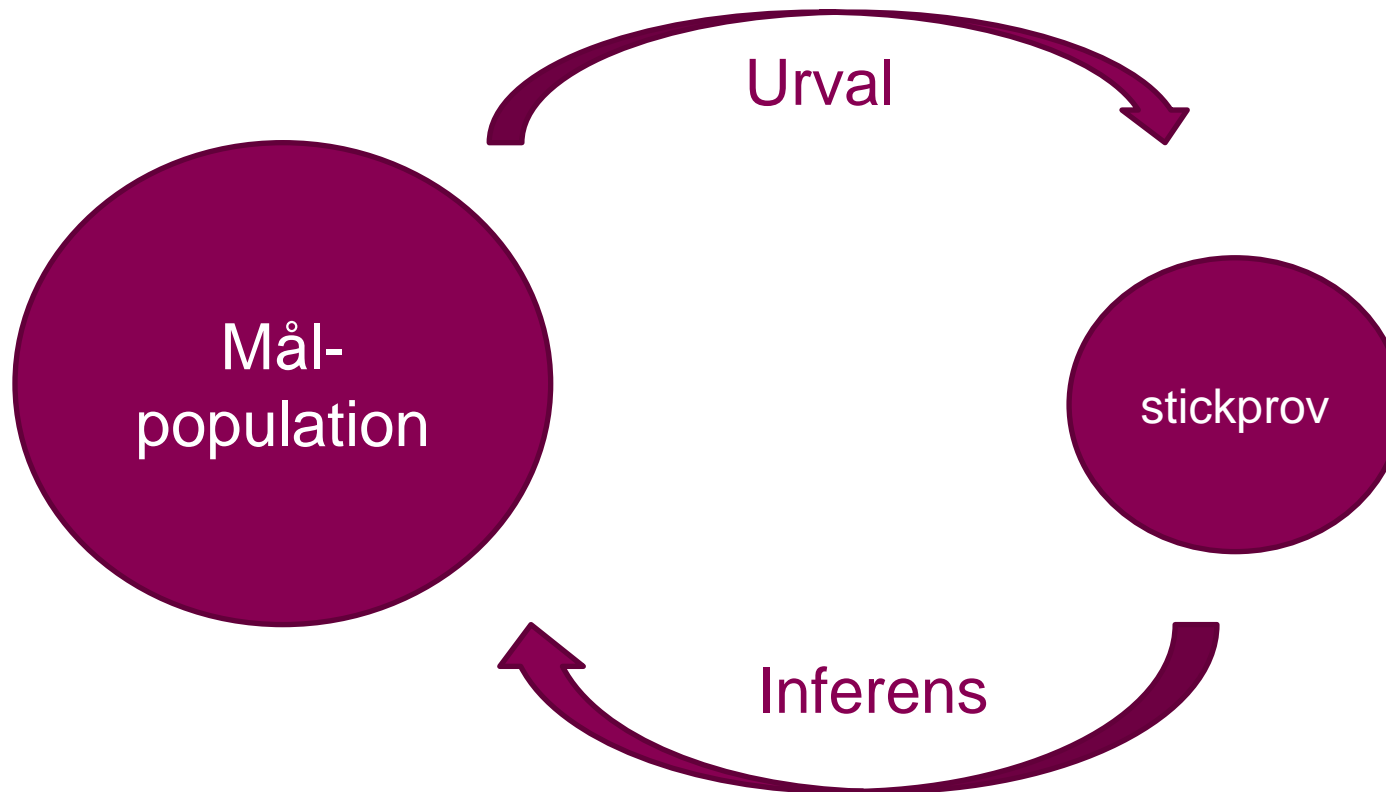
kräver fler antaganden än de som ingår i beräkningar av t ex konfidensintervall.

Hur bör totalundersökningar redovisas?

Hur ska då totalundersökningar redovisas? Har man följt hela den svenska befolkningen under en viss period så är det som kommer fram korrekt med reservation för eventuella brister i de register som använts och de beräkningar som gjorts. Vill man sedan generalisera sina

Grundläggande statistikteori

– stickprovsundersökning



Stickprovsundersökning

- *”Vid en stickprovsundersökning undersöks en del av populationen. Stickprovsundersökningar har stor praktisk användning och begagnas t.ex. vid opinionsmätningar, försöksverksamhet och statistisk kvalitetskontroll.”*
- *”En stickprovsundersökning ger i regel ett osäkrare resultat än en totalundersökning.”*

Blom och Holmqvist,

Statistikteori med tillämpningar, Studentlitteratur

Totalundersökning

- *”I en totalundersökning studeras hela populationen. Totalundersökningar har stor betydelse på många områden, t.ex. inom offentlig statistikproduktion och inom industrin. Ofta är en totalundersökning alltför dyrbar eller tidsödande att genomföra, ibland principiellt eller praktiskt otänkbar. Totalundersökningar kräver inte någon statistikteori och behandlas nästan inte alls i fortsättningen”*

Blom och Holmqvist,

Statistikteori med tillämpningar, Studentlitteratur

Läkartidningen 2004

Cancerforskare, sluta redovisa konfidensintervall när det inte behövs!

II Med förvåning kan man i vetenskapliga tidskrifter läsa svenska artiklar som redovisar osäkerhet i form av konfidensintervall eller p-värden trots att det av metodbeskrivningen framgår att undersökningen är en populationsbaserad totalundersökning som använt det svenska cancerregistret, ofta i kombination med andra heltäckande register som t ex registren över slutenvård eller folk- och bostadsräkningen.

kräver fler antaganden än de som ingår i beräkningar av t ex konfidensintervall.

Hur bör totalundersökningar redovisas?

Hur ska då totalundersökningar redovisas? Har man följt hela den svenska befolkningen under en viss period så är det som kommer fram korrekt med reservation för eventuella brister i de register som använts och de beräkningar som gjorts. Vill man sedan generalisera sina

- *Med förvåning kan man i vetenskapliga tidskrifter läsa svenska artiklar som redovisar osäkerhet i form av konfidensintervall eller p-värden trots att det av metodbeskrivningen framgår att undersökningen är en populationsbaserad totalundersökning*
- *Det finns ingen statistisk osäkerhet till följd av urval och därmed inget behov av konfidensintervall*

Bakomliggande argument

Ändlighetskorrektion:

Om vi har en ändlig population och tar ett stickprov som är mer än ca. 10 % av populationen, dvs. urvalskvoten $\frac{n}{N} > 0,1$, så används ändlighetskorrektion. Vi får en säkrare skattning och korrigerar därför medelfelet så att

det minskar med faktorn $\sqrt{\frac{N-n}{N-1}} \approx \sqrt{1-\frac{n}{N}}$

Ex på ändlighetskorrektion:

KI för μ :

$$\bar{x} \pm z \cdot \frac{s_x}{\sqrt{n}} \cdot \sqrt{1-\frac{n}{N}} = \bar{x} \pm z \cdot \sqrt{\frac{s_x^2}{n} \left(1-\frac{n}{N}\right)}$$

□

Slutsats

“När vi nu har ett cancerregister som ger facit till en mängd intressanta frågeställningar så borde vi i Sverige kunna ta forskningen till en högre nivå än att nöja oss med att redovisa konfidensintervall, vilka alltså inte ger någon extra kunskap.”

Reaktion på artikeln i Läkartidningen

Korrespondens

Cancerforskare behöver statistisk osäkerhet!

II Henrik Dal och Mikael Andersson (D/A) uppmanade i Läkartidningen 4/2004 (sidan 314) cancerforskarna att låta bli att ange konfidensintervall eller osäkerhet vid studier som utnyttjar svenska populationsbaserade register.

talet bröstcancerfall år 2002. Tar man ett medelvärde av de möjliga utfallen skulle det återspegla den underliggande bröstcancerfrekvensen, och därmed också den process som ger upphov till fallen.

den u
förme
t ex sl
frekve
oss at
under

*Paul Dickman, Juni Palmgren, Yudi Pawitan
Läkartidningen, Nr 20, 2004*

Huvudresonemang

- Ett komplett urval bör för att besvara vetenskapliga frågeställningar ses som ett uttryck (realisation) för en underliggande slumpmässig process.
- Det vi observerar innehåller både systematik och slump, och det är den underliggande systematiken som förmedlar det vetenskapliga budskapet, t ex skillnader i underliggande cancerfrekvens mellan olika grupper.

Exempel

År 2002 inträffade 95 fall av tungcancer bland svenska män, medan det bland kvinnor bara förekom 70 fall. En jämförelse av de observerade värdena 95 och 70 har inget vetenskapligt innehåll. Däremot är det meningsfullt att fråga sig om den underliggande frekvensen av tungcancer är olika för män och kvinnor. Även om värdena 95 och 70 kommer från ett komplett urval är de intressanta först då de behandlas som ett utfall från en process som innehåller slump.

Sammanfattningsvis

- Beroende vad vi är intresserade av för fråga så kan det av olika anledningar vara relevant att mäta osäkerhet i utfallet (även om vi har alla data tillgängliga)
- Slumpvariabel $X \sim \text{fördelning}(\theta)$, vi är intresserade av θ inte X

Journal of the American Statistical Association 1941;36:45-49

ON THE INTERPRETATION OF CENSUSES AS SAMPLES

BY W. EDWARDS DEMING AND FREDERICK F. STEPHAN
Bureau of the Census

As a basis for scientific generalizations and decisions for action, a census is only a sample. In addition to serving the function of an inventory as of a certain date, the census tabulations serve also another important objective, namely, as bases for prediction. Any social or economic generalization, and any recommendation for a course of action, involves a prediction. For such purposes, the census takes on the character of a sample.

Journal of the American Statistical Association 1953;48:244-255

ON THE DISTINCTION BETWEEN ENUMERATIVE AND ANALYTIC SURVEYS*

W. EDWARDS DEMING

Bureau of the Budget and New York University

*The distinction between the enumerative and analytic uses of data.¹ Briefly, the enumerative question is how many? The analytic question is *why?* is there any difference between the two classes, and if so, how big are the differences?*

Klassifikation av studier (utifrån syfte)

Inventering (uppräkning)

- Sammanställning av data som besvarar frågan:

Hur många?

- Rapporterade cancerfall - planering av vårdplatser
- Barnkullar - planering av dagis/skol platser

Vetenskaplig analys

- Undersökning som syftar till att besvara frågan:

Varför?

- Systematiska skillnader mellan grupper eller över tid? Faktorer som påverkar de drivande processerna?

Att tolka en totalundersökning som ett stickprov

- *”A census describes a population that is subject to the variations of chance, because it is only one of the many possible populations that may have resulted from the same underlying system of social and economic causes”*

Totalundersökning – ett stickprov ur en superpopulation

- *”From the point of view being expressed here, however, even a complete census, for scientific generalizations, describes a population that is but one of the infinity of populations that will result by chance from the same underlying social and economic cause system. This infinity of populations may itself be thought of as a **superpopulation**. A sample enquiry is then a sample of a sample, and a so-called 100 per cent sample is simply a larger sample, but is still only a sample.”*

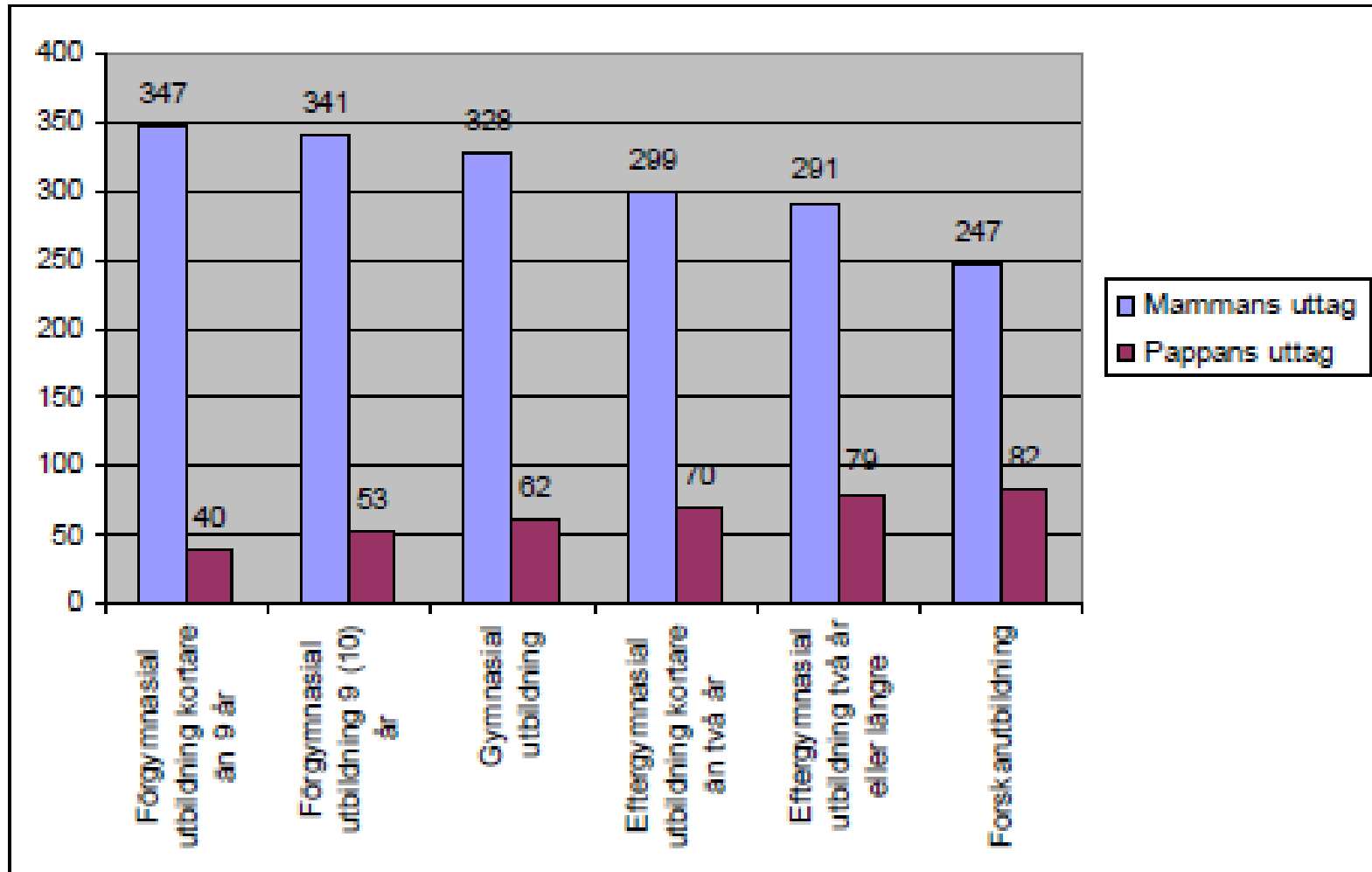
Föräldrapenning

Båda föräldrarnas försäkring?

Socialförsäkringsrapport 2011:13

- I denna rapport **redovisar** och **analyserar** Försäkringskassan utvecklingen inom föräldrapenningen med inriktning på fördelning av föräldrarnas uttag av föräldrapenning.
- I kapitel 3 analyseras hur utfallet av olika föräldrapenningmått varierar beroende på föräldrarnas kön, ålder, utbildning och födelseregion. I det fjärde kapitlet studeras effekterna av de senaste årens reformer med fokus på hur dessa påverkat jämställdheten i föräldrapenninguttaget.

Figur 20 Uttag av nettodagar uppdelat på föräldrarnas utbildning



Statistica Sinica 8(1998), 1131-1151

VARIANCE ESTIMATION FOR SUPERPOPULATION PARAMETERS

Edward L. Korn and Barry I. Graubard

Abstract: In scientific applications, interest usually focuses on the “superpopulation” parameters of a stochastic model hypothesized to underlie the generation of the values in a finite population, rather than finite-population parameters themselves. Variance formulas for sampled data that incorporate finite-population correction factors are not appropriate for these applications. For simple random sampling, it is common practice to ignore these correction factors in variance estimation; this yields correct superpopulation inference under a simple superpopulation model.

Möjliga förklaringar till ett observerat samband mellan exponering och utfall i observationella studier

- Bias (exempelvis selektionsbias, re-call bias etc)
- Confounding (egentligen en form av bias)
- Slump
- En verklig association mellan exponering och utfall

Läärä (2011) in *Methods in Biobanking*

The estimation of a parameter is prone to error; we can express an estimate as a sum of three components:

Estimate = true parameter value + bias + random error.

Schneeweiss (7). The main sources of *random error* are in turn (a) biological variation between and within individuals, (b) measurement variation, (c) sampling (whether random or non-random), and (d) division of exposure (whether properly randomized or non-randomized).

Borde vi presentera mått på osäkerhet då vi känner till hela populationen?

- JA – om syftet med den statistiska undersökningen är vetenskapligt (analytiskt).
- Om syftet istället är administrativt finns inga skäl till att redovisa standardavvikelser, konfidensintervall etc.
- Oklarheten består inte i om osäkerhet ska redovisas utan hur den i själva verket ska mätas – dvs hur bör man härleda variansformler då målpopulationen är en högst hypotetisk superpopulation?
- Tillämpar i dagsläget ett frekventistiskt synsätt.

Förslag till diskussionspunkter

- Vanliga missuppfattningar (generaliserbarhet till andra populationer, prediktionssyfte)
- Superpopulationen är inte väldefinierad, så vad kan vi egentligen säga om den ”enda observation” som vi har att tillgå?
- Frågor som ni får på försäkringskassan angående detta?
- Övriga synpunkter?

Vill du läsa mer?

- Dahl H, Andersson M; *Cancerforskare, sluta redovisa konfidensintervall när de inte behövs!* (2004), Läkartidningen, Nr 4, Vol 101
- Dickman P, Pawitan Y, Palmgren J; *Cancerforskare behöver statistisk osäkerhet* (2004), Läkartidningen, Nr 20, Vol 101
- Ralph Nilsson; *Till Konfidensintervallens försvar* (2004), Läkartidningen, Nr 20, Vol 101
- Andrew Gelmans blogg (<http://andrewgelman.com/>)
 - How does statistical analysis differ when analyzing the entire population rather than a sample? (2009)
 - How do you interpret standard errors from a regression fit to the entire population? (2011)