



Krav på stickprovsstorlek vid inventering av bebyggelsens avvikelser från originalskick

Rapportering av ett uppdrag från Riksantikvarieämbetet

Ulf Grandin

Department of Environmental Assessment
Swedish University of Agricultural Sciences
Box 7050, SE 750 07 Uppsala

Krav på stickprovsstorlek vid inventering av bebyggelsens avvikelser från originalskick

Rapportering av ett uppdrag från Riksantikvarieämbetet

Ulf Grandin

ISSN 1403-977X

Inledning

Hur många byggnader måste man inventera för att ha en möjlighet att kunna upptäcka en signifikant förändring från originalskick på bebyggelse på en ort? Syftet med denna rapport är att lämna ett underlag till svaret på den inledande frågan. Som det kommer att framgå går det inte att lämna ett entydigt svar då frågan ger upphov till en rad ställningstaganden som måste klaras ut innan det går att använda statistik för att beräkna svaret. Nya frågor som dyker upp är bl.a.:

- hur ska man värdera byggnaders förändring från originalskick,
- hur stor förändring måste till innan förändringen är noterbar ur ett antikvariskt/kulturmiljöperspektiv,
- vilken risk är man villig att ta att förändringen p.g.a. slumpen bara beror på ett icke-representativt stickprov,
- hur stor chans att ha rätt vill man ha då man påstår att det skett en förändring?

Bedömning av byggnader

För att på ett förhållandevis snabbt och enkelt vis kunna göra mätningar av förändringar av byggnader har det utvecklats en kvantitativ mätmetod av byggnaders avvikelse från originalskick (se t.ex. Katzeff 2007). Metoden går ut på att poängsätta byggnader, byggnadsdel för byggnadsdel, utgående från hur välbevarade byggnaderna är i förhållande till ursprunget. Varje byggnadsdel bedöms i en tregradig skala där 2 poäng motsvarar en intakt ursprunglig byggnadsdel, 1 poäng en förändrad men anpassad byggnadsdel och 0 poäng en förvanskad byggnadsdel. De byggnadsdelar som bedömts är fasad, tak, fönster, dörrar och balkong.

Styrkeanalyser och hypotesprövning

En vanlig fråga från personer som bedriver miljöövervakning är hur många stickprover som behövs. Detta är en fråga man ställer innan man startar ett övervakningsprogram. Svaret på denna fråga beror av fem olika faktorer, och sammanfattas inom en familj av statistiska metoder som kallas styrkeanalys. Efter att data är insamlade kan man t.ex. beräkna om det skett en förändring med metoder som bygger på begreppet statistisk hypotesprövning. Syftet med denna rapport är dock att beräkna antalet prover som behövs för att kunna upptäcka förändringar i bebyggelse enligt det system som Katzeff (2007) använder för att skatta förändringar. Som en bakgrund kommer dock även principen för att beräkna om det skett en förändring presenteras.

Risk och chans

Inom statistisk hypotesprövning får man svar på frågan: ”Hur stor är risken att jag har fel när jag påstår att det skett en förändring?”, eller liknande frågeställningar. Om ett statistiskt test visar att det är mindre än 5 % risk, brukar man dra slutsatsen att de data man samlat in pekar på att det skett en förändring. Man kan också fråga sig hur stor chans det är att man har rätt i de fall man påstår att det skett en förändring. Intuitivt kan man tycka att denna chans borde vara 95 %, om risken att ha fel är 5 %. Nu är det dock inte så enkelt! Risken att ha fel är kopplad till chansen att ha rätt på ett mer intrikat sätt än så (se t.ex. Grandin 2003b). Inom styrkeanalys måste man själv välja nivån på dessa faktorer. Vanligtvis väljer man högst 5 % för risken att ha fel och mellan 80 % och 95 % för chansen att ha rätt.

Har det skett en förändring?

Svaret på frågan i rubriken kräver i de flest fall en motfråga: vad menas med en förändring. Man kommer alltid att hitta en förändring, men vad är en förändring som har någon relevans? En förändring från t.ex. 1 565 473 till 1 565 475 stycken lövsångare är visserligen en förändring men ur biologisk synpunkt har det ingen relevans. Dessutom är ökningen på 2 individer förmodligen inom felräkningsmarginalen. Motsvarande resonemang gäller givetvis även för förändringar i bebyggelse.

Eftersom övervakning bygger på stickprov finns det alltid en risk att stickproven utgör extremer bland de objekt man har att välja sitt stickprov från. Av detta följer att man med små stickprov har större risk att råka få ett skevt och icke-representativt stickprov. En förlängning av detta är att ju mindre förändring man vill kunna upptäcka desto större stickprov krävs för att kunna skilja en liten förändring från slumpmässiga skillnader som uppkommer på grund av att stickprovet är litet.

Olika typer av jämförelser mellan då och nu

Ett övervakningsprogram bygger i princip på en jämförelse av situationen före och efter, eller då och nu. Visserligen kan program sträcka sig över fler än två observationstillfällen, men den grundläggande principen är ändå en jämförelse mellan två observationer.

Två observationstillfällen kan vara sammankopplade på två sätt. Antingen observerar man samma objekt vid två tillfällen, eller så observerar man slumpvis valda objekt inom ett område vid två tillfällen. Båda metoderna har sina fördelar och nackdelar. Om man vill uttala sig om förändring i byggnadsstil i allmänhet inom en stad eller annat avgränsat område samlar man data från slumpvis valda byggnader vid ett tillfälle, och nya slumpvis valda byggnader vid ett annat tillfälle. Denna typ av tester brukar kallas operade. Är man däremot intresserad av att följa specifika hus inom en stad samlar man data från slumpvis valda byggnader och återkommer till samma byggnader vid nästa observationstillfälle, s.k. parade tester.

Ytterligare ett sätt att jämföra medelvärden är att relatera medelvärdet från en inventering till ett referensvärde. I denna studie är detta tillämpligt för att undersöka om det skett en förändring från originalskick till observerat skick.

Metoder

Normering av poängskala

I det datamaterial som ligger till grund för beräkningarna i denna rapport var husen oftast poängsatta i en skala där högsta poängsumma är 10. I några fall var den högsta poängsumman endast 8 och i andra fall kunde summan uppgå till 12. För att kunna använda alla observationer har poängsumman i de fall maxpoängen var annat än 10 normerats enligt:

$$\text{Normerad poäng} = \frac{\text{poäng} \times 10}{\text{max poäng}}. \quad (\text{Ekvation 1})$$

Åldersindelning av hus

För att få ett mer homogent dataunderlag, och för att underlätta det praktiska arbetet har underlagsdata delats upp i åldersklasser med avseende på husens ursprung. Hus byggda efter 1930 har delats in i decennieklasser, medan hus äldre än från 1930 slagits samman till en klass.

Hur beräknar man erforderlig antal prover

För att beräkna erforderligt antal prover krävs data på det man vill undersöka. Utöver data krävs det även att man själv måste ange värdet på ett antal variabler. Dessa variabler är:

- den risk man är villig att ta för att riskera att en observerad förändring endast föreligger i stickprovet och inte i hela den population man undersöker (signifikansnivån),

- den chans man vill ha att man har rätt då man påstår att det skett en förändring (den statistiska styrkan),
- åt vilket håll en förändring kan gå,
- hur stor förändring man vill kunna upptäcka.

Risken i punkt 1 har i denna studie satts till standardnivån 5 %. Likaså har chansen i punkt 2 satts till standardnivån 80 %. Riktning på eventuella förändringar är satta så att det bara kan ske en förändring åt ett håll. Antingen är ett hus i originalskick eller så är det omgjort, och får då en lägre poäng. Dessutom går det inte att få högre poäng än originalskick, och därför kan en eventuell förändring bara ske i en riktning vid den typ av inventering som förutsätts i denna rapport. I statistiska termer kallas detta för att man gör ensidiga, eller ensvansade, tester. Teoretiskt är det givetvis fullt möjligt att ett hus blir restaurerat till originalskick. En sådan förändring kan dock bara upptäckas i de fall man återkommer till samma hus vid en upprepad inventering. Den här rapporten bygger dock på att man inte återinventerar samma hus, utan istället gör ett nytt slumpmässigt urval av hus vid en upprepad inventering eller jämför resultaten från en första inventering mot originalskick. I det senare fallet är det omöjligt att notera en restaurering i den skala som används för poängsättning. Storleken på den förändring som man vill kunna upptäcka går inte att beräkna med något statistiskt test. I denna studie kommer därför erforderligt antal hus för förändringar mellan 0.5 och 5 poäng att presenteras.

De data man måste ha för att beräkna erforderligt antal prover är mått på variationen hos den variabel man studerar. I detta fall är det variationen i poängsumman som är av intresse. Detta är beräknat som standardavvikelsen för den normerade poängsumman, för respektive åldersklass och för hela materialet.

För att beräkna lägsta erforderliga antal hus som måste poängsättas för att kunna detektera en signifikant förändring har följande ekvation använts:

$$n \geq \frac{\sigma^2 (z_{1-\alpha} + z_{1-\beta})^2}{MDS^2}, \text{ där} \quad (\text{Ekvation 2})$$

n = antal hus som måste undersökas i inventering nummer 2 vid ett omdrev,

σ = standardavvikelsen för det material som studien baseras på,

α = vald signifikansnivå, här satt till 0,05,

$1-\beta$ = vald statistisk styrka, här satt till 0,8,

$z_{1-\alpha}$ = z-koefficienten för $1-\alpha$,

$z_{1-\beta}$ = z-koefficienten för den valda styrkan,

MDS = Minsta Detekterbara Skillnad. Anges i absoluta tal. En förändring på t.ex. 20 % från originalskick (poäng = 10) anges som $0,2 \times 10 = 2$.

Resultatet av denna ekvation avrundas alltid till närmsta högre heltal. Ett resultat på t.ex. 20,02 visar att det behövs fler än 20 hus. Eftersom antal hus är en heltalsvariabel måste detta tal avrundas till 21.

Vid ett omdrevsförfarande bör man sträva efter att ha lika många observationer vid båda inventeringarna. Ekvation 2 ger dock antalet hus som krävs vid inventering nummer 2 i ett omdrev.

Hur beräknar man om det skett en signifikant förändring

Syftet med denna studie är att undersöka hur många hus som måste inventeras för att hitta en eventuell förändring. Det är dock inte säkert att det föreligger någon förändring. I detta avsnitt presenteras kort olika varianter av den statistiska metod som används för att undersöka om det skett någon förändring, s.k. t-tester. För en mer ingående beskrivning hänvisas till t.ex. Grandin (2003a).

Ett statistiskt test som prövar om det skett en förändring är ett verktyg för att svara på om den förändring man ser i ett stickprov kan generaliseras till hela den population man undersöker, eller om det finns en risk att stickprovet råkat bli en samling av extremvärden. Denna risk är den s.k. signifikansnivån. Vanligtvis sätter man den högsta gränsen för denna risk till 5 %.

De beräkningar som är utförda för att skatta erforderligt antal byggnader att poängsätta är baserade på så kallade oparade tester, d.v.s. oberoende slumpvisa stickprov. Skattingen av erforderligt antal hus kan antingen användas för att dimensionera ett omdrev eller för att undersöka om det skett en signifikant förändring jämfört med originalskick. I båda fallen beräknar man en s.k. testvariabel som i denna typ av tester kallas t. Det framräknade värdet på t måste sedan jämföras med gränsvärden som finns tabellerade i de flesta statistikböcker (t.ex. Sokal och Rohlf 1995). I en sådan tabell jämför man det framräknade värdet med det gränsvärde som är tabellerat för den valda signifikansnivån och för aktuellt antal frihetsgrader. Om det framräknade värdet överstiger gränsvärdet kan man dra slutsatsen att det skett en förändring i huspoäng mellan de båda undersökningstillfällena.

För att undersöka om det skett en signifikant förändring jämfört med originalskick (poängsumma 10) används följande formel:

$$t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)}, \text{ där} \quad (\text{Ekvation 3})$$

\bar{x} = det medelvärde man erhållit efter en inventering,
 μ = det teoretiska medelvärdet, i detta fall maxpoängen (= 10),
 s = standardavvikelsen i de data man samlat in, och
 n = antalet inventerade hus.

Det framräknade värdet jämförs sedan med tabellerade data för n-1 frihetsgrader.

För att undersöka om det skett en förändring mellan två omdrev används istället följande formel:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE_{(1-2)}}, \text{ där} \quad (\text{Ekvation 4})$$

\bar{x}_1 och \bar{x}_2 är medelvärdena av huspoängen vid undersökning 1 respektive 2,
 $SE_{(1-2)}$ är medelfelet för skillnaden mellan medelvärdena av undersökning 1 respektive 2, och beräknas enligt:

$$SE_{(1-2)} = \sqrt{\frac{s_{(1,2)}^2}{n_1} + \frac{s_{(1,2)}^2}{n_2}}, \text{ där} \quad (\text{Ekvation 5})$$

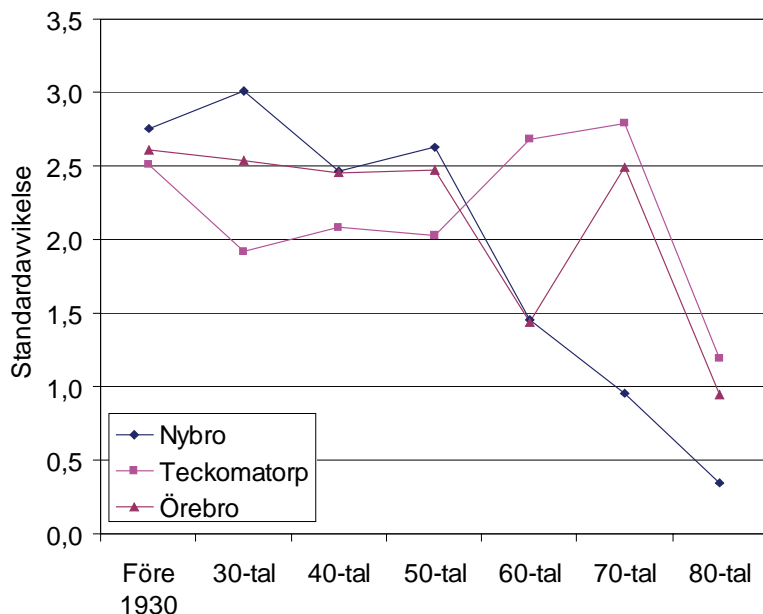
$$s_{(1,2)}^2 = \frac{(n_1 - 1) \times s_1^2 + (n_2 - 1) \times s_2^2}{(n_1 - 1) + (n_2 - 1)}, \text{ där} \quad (\text{Ekvation 6})$$

s_1^2 och s_2^2 är de olika provens varianser, och
 n_x är antal observationer i respektive grupp.

Det framräknade värdet jämförs sedan med tabellerade data för $(n_1 + n_2 - 2)$ frihetsgrader.

Resultat

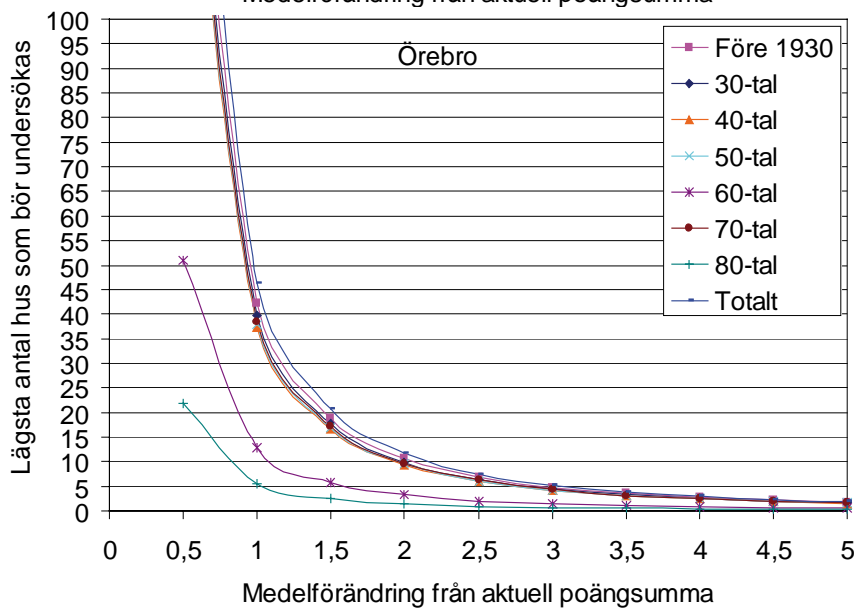
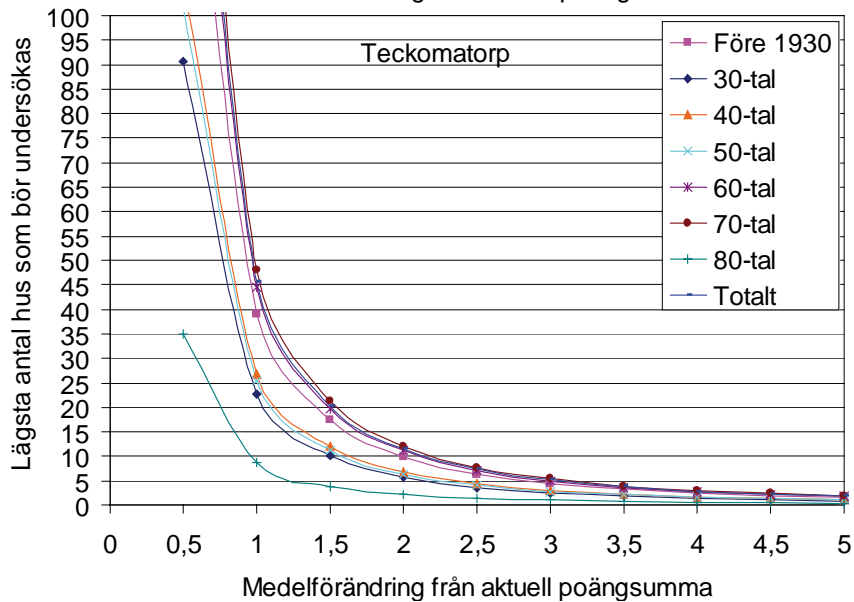
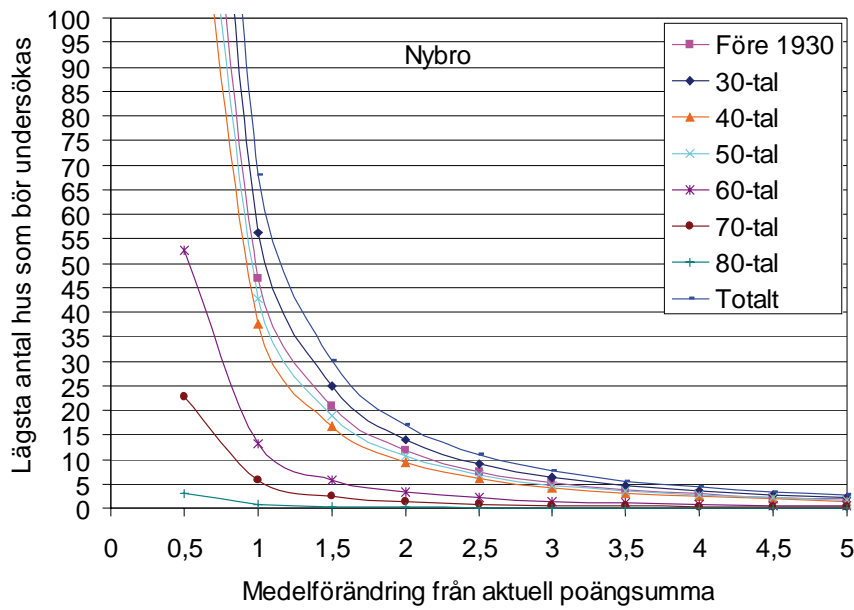
För att beräkna erforderligt antal hus att inventera krävs data på variationen i aktuell status. För denna rapport fanns data från tre orter; Nybro, Teckomatorp och Örebro. Det kan konstateras att variationen är likartad för hus uppförda före 1960, och att det är låg variation i avvikelse från originalskick bland hus uppförda under 1980-talet (Fig. 1 och Bilaga, tabell 1). För hus uppförda före 1930 var variationen så likartad för de olika decennierna att dessa hus kunde slås samman till en grupp.



Figur 1. Standardavvikelse i byggnadspoäng för hus uppförda under olika tidpunkter, uppdelat på tre orter.

Utifrån dessa data på variation och de antaganden som är presenterade i metodavsnittet har erforderligt antal hus för att kunna upptäcka en förändring beräknats. Resultaten presenteras dels i diagramform (Fig. 2), dels som tabeller i en bilaga (Bilaga, tabell 2). Diagrammen kan användas både för att undersöka minsta stickprovsstorlek för att undersöka om det föreligger en signifikant förändring jämfört med originalskick, samt för att undersöka det antal hus som behövs i en omdrevsinventering. Vid en jämförelse med originalskick visar diagrammen det antal hus som lägst måste inventeras. Om undersökningen gäller ett omdrev visar diagrammen lägsta antal hus i respektive undersökning. Det behövs t.ex. 6 hus i vardera undersökningen om man vill kunna detektera en medelförändring på 1 poängenheter hos hus uppförda under 1970-talet i Nybro (Figur 2, översta diagrammet, kurvan för 1970-tal).

Tack vare en relativt låg variation blir kraven på antal hus riktigt höga först när man vill kunna upptäcka medelförändringar lägre än ca 1 poängenheter. Detta resultat upprepas i stort sett för alla undersökta orter, och åldrar på hus. Utan uppdelning i åldersklasser blir variationen högre, vilket medför krav på fler inventerade hus för att kunna upptäcka en eventuell förändring. För exemplet med hus i Nybro ovan behövs det istället för 6 hus uppförda på 1970-talet istället minst 69 hus om man inte tar hänsyn till byggnadsår (Figur 2, översta diagrammet, kurvan för Totalt). Utan uppdelning på vare sig ort eller byggnadsår blir standardavvikelsen 3,04, vilket närmast motsvaras av kurvan för hus från 1930-talet i Nybro, i figur 2.



Figur 2. Diagram som visar lägsta antal hus som måste undersökas för att i oparade t-tester kunna upptäcka olika stora förändringar i den skala som används för att klassa hus, uppdelat på ort och byggnadsår.

Diskussion och kommentarer

Resultaten i denna studie är helt baserade på medelvärden av avvikelser från originalskick. Hur man delar in ett datamaterial i undergrupper för att beräkna medelvärden och tillhörande mått på variation, beror på frågeställning. I den här studien är medelvärdena grupperade efter ort och byggnadsår. I strikt mening gäller resultaten bara de orter som stått för underlagsdata. Extrapolering av resultaten till andra orter och grupperingar av byggnadsår ska därför ske med vetskap om att resultaten i denna rapport då endast blir vägledande. Genom att utföra en pilotstudie i ett nytt område kan variationen beräknas och jämföras med de värden på variation som ligger till grund för denna rapport. Om variationen i ett nytt område är av samma storleksordning som här, kan resultaten översättas.

Resultaten i denna studie är uppdelat på olika byggnadsår. Här är det viktigt att komma ihåg att hus i framtiden inte kommer att se ut som vid inventeringstillfället. Därför gäller kurvorna för de olika decennierna egentligen bara vid inventeringstillfället. I framtiden kommer husen från 1980-talet att variera mycket mer än idag och därmed ökar också det erforderliga antalet hus som krävs för att kunna detektera en förändring. I princip skulle hälften av de inventerade husen från 1980-talet kunna ha blivit föremål för omfattande renoveringar dagen efter att inventeringen utfördes, och därmed skulle variationsmåttet för det decenniet var helt felaktigt. En grov generalisering skulle dock kunna vara att för 1980-talshusen om 20 år använda resultaten för husen från 60-talet, men det är inte säkert att utvecklingen är så rätlinjig.

Hittills har hela denna rapport behandlat undersökningar som bygger på s.k. operade data, och att en förändring bara kan upptäckas åt ett håll. I de fall man utför en återinventering på samma hus erhåller man istället s.k. parade data. Vid en sådan inventering är det teoretiskt möjligt med förändringar åt två håll. Detta medför att kurvorna i figur 2 inte längre kan användas. Detsamma gäller om man utför operade undersökningar där man kan kvantifiera förändringar åt två håll. I dessa fall kan dock resultaten från figur 2 konverteras till att gälla för tvåsidiga tester genom att multiplicera lägsta erforderligt antal hus med 1,3. Egentligen är detta mer komplicerat, men denna förenklade omvandling ger en tillräckligt god precision.

Metoden för att beräkna om det verkligen skett en förändring blir även den annorlunda i de fall man har parade data. I ekvation 3 ersätts \bar{x} med medelvärdet av alla differenser mellan resultaten från inventering 1 och 2, och μ sätts till 0.

Referenser

Grandin, U (2003a). *Dataanalys och hypotesprövning för statistikanvändare*. Naturvårdsverket. Stockholm.

http://www.naturvardsverket.se/upload/02_tillstandet_i_miljon/Miljoovervakning/handledning/dataanalys_och_hypotesprovn.pdf

Grandin, U. (2003b). *Kravet på noggrannhet i vattendirektivet. Några exempel på beräkningar*. Naturvårdsverket. Stockholm.

http://www.naturvardsverket.se/upload/02_tillstandet_i_miljon/Miljoovervakning/handledning/kravet_pa_noggrannhet.pdf

Katzeff, A (2007). *Miljömålsprojektet "Operation Kungsörn". Utvärdering och utveckling av en indikator för mätning av bebyggelsemiljöers förändringar*. Rapport länsstyrelserna i Kalmar län, Skåne län, Västra Götalands län och Örebro län. Länsstyrelsen i Skåne Län, Miljöavdelningen. Malmö.

http://www.m.lst.se/m/Publikationer/Rapporter/2007/Operation_Kungsorn.htm

Sokal, R.R., Rohlf, F.J., (1995). *Biometry: The Principles and Practice of Statistics in Biological Research*. Freeman, New York.

Bilaga

Tabell 1. Variation i husbedömningspoäng, uppdelat på byggnadsår och ort. Ju högre värde desto större variation.

Byggnadsår	Standardavvikelse		
	Nybro	Teckomatorp	Örebro
Före 1930	2,75	2,51	2,61
30-talet	3,01	1,92	2,53
40-talet	2,46	2,08	2,46
50-talet	2,63	2,03	2,47
60-talet	1,46	2,68	1,44
70-talet	0,96	2,79	2,49
80-talet	0,35	1,19	0,94
Totalt	3,32	2,72	2,74

Tabell 2. Lägsta antal hus som måste undersökas för att kunna upptäcka olika stora förändringar i den skala som används för att klassa hus, uppdelat på hus från olika tidsperioder. Alla siffror avrundade till närmast högre heltal.

Förändring från aktuell poängsumma	Nybro							
	Före 1930	30-tal	40-tal	50-tal	60-tal	70-tal	80-tal	Totalt
0,5	188	225	150	171	53	23	3	273
1	47	57	38	43	14	6	1	69
1,5	21	25	17	19	6	3	1	31
2	12	15	10	11	4	2	1	18
2,5	8	9	6	7	3	1	1	11
3	6	7	5	5	2	1	1	8
3,5	4	5	4	4	2	1	1	6
4	3	4	3	3	1	1	1	5
4,5	3	3	2	3	1	1	1	4
5	2	3	2	2	1	1	1	3

Förändring från aktuell poängsumma	Teckomatorp							
	Före 1930	30-tal	40-tal	50-tal	60-tal	70-tal	80-tal	Totalt
0,5	157	91	108	102	178	193	35	183
1	40	23	27	26	45	49	9	46
1,5	18	11	12	12	20	22	4	21
2	10	6	7	7	12	13	3	12
2,5	7	4	5	5	8	8	2	8
3	5	3	3	3	5	6	1	6
3,5	4	2	3	3	4	4	1	4
4	3	2	2	2	3	4	1	3
4,5	2	2	2	2	3	3	1	3
5	2	1	2	2	2	2	1	2

Förändring från aktuell poängsumma	Örebro							
	Före 1930	30-tal	40-tal	50-tal	60-tal	70-tal	80-tal	Totalt
0,5	169	159	150	152	52	154	22	186
1	43	40	38	38	13	39	6	47
1,5	19	18	17	17	6	18	3	21
2	11	10	10	10	4	10	2	12
2,5	7	7	6	7	3	7	1	8
3	5	5	5	5	2	5	1	6
3,5	4	4	4	4	2	4	1	4
4	3	3	3	3	1	3	1	3
4,5	3	2	2	2	1	2	1	3
5	2	2	2	2	1	2	1	2