

**Inlämningsuppgift i Undersökningsmetodik, 3 hp,  
ingående i kursen Regressionsanalys och undersökningsmetodik, 15 hp.**

**Del 1: Olika urvalsmetoder – punkt- och intervallskattningar**

Enligt kursbeskrivningen gäller att  
”Efter att ha genomgått kursen förväntas studenten kunna  
...  
argumentera för och genomföra olika slumpmässiga urval  
...”

Denna uppgift utföres individuellt och redovisas skriftligt. Särskilt svarsblad ska användas!  
Sista inlämningsdatum mm: vänligen se Momentbeskrivningen för aktuell termin!  
Glöm inte att skriva namn och gruppnummer på ditt inlämnade arbete!

På denna uppgift ges något av betygen godkänd eller underkänd. Uppgiften kan behöva kompletteras för att uppnå betyget godkänd. Endast ett tillfälle ges för komplettering.

*Målsättningen med denna uppgift är att genomföra tre olika slumpmässiga urval av ett antal kommuner från populationen *samtliga kommuner i landet*. De olika urvalen skall ske genom obundet slumpmässigt urval, stratifierat urval med obundet slumpmässigt urval inom strata samt genom systematiskt urval. För varje urval skall en punktskattning (och i vissa fall ett konfidensintervall) bestämmas för parametern *genomsnittlig kommunalskatt 2007 i landets samtliga kommuner*.*

Varje student erhåller en specifik sida med slumptal, med vars hjälp man skall bestämma vilka kommuner som skall ingå i urvalen. Dessa är här skapade med hjälp av Minitabs slumpgenerator.

För var och en av kommunerna i populationen finns det värden på sex olika bakgrundsvariabler samt värden på undersökningsvariabeln, dvs på den totala skattesatsen i kommunen 2007. Värdena på skattesatsen skall här betraktas som kända *enbart* för de kommuner som väljs i de olika urvalen.

Materialet är hämtat från Årsbok för Sveriges kommuner samt från SCB:s databaser. De olika variablerna är

- 1 Kommunens namn
- 2 Länsnummer
- 3 Folkmängd den 31 mars 2007
- 4 Medelinkomst 2005 (tusen kr), beräknad som ”Sammanräknad förvärvsinkomst (inkomst av tjänst och inkomst av näringsverksamhet) för befolkningen 20 år och däröver”.
- 5 Andel (%) förvärvsarbetande kvinnor inom den offentliga sektorn år 2000.
- 6 Andel (%) moderata kommunfullmäktigeledamöter efter valet 2006.
- 7 Total skattesats (kronor per 100 skattekronor) i kommunen 2007.

Materialet finns återgivet i en bilaga och är sorterat på fem olika sätt, nämligen i stigande storleksordning efter

- länsnr
- folkmängd
- medelinkomst
- procentandel kvinnor inom den offentliga sektorn
- procentandel moderata ledamöter i kommunfullmäktige

Som en lämplig ram vid de olika urvalsförfarandena kan kommunerna, sorterade efter en eller flera av dessa principer, med fördel användas.

## Uppgifter:

I samtliga urvalsförfaranden är den sökta parametern ”**Den genomsnittliga kommunalskatten 2007 i landets samtliga kommuner**”.

1. Gör ett *obundet slumpmässigt urval utan återläggning* av 30 kommuner samt bestäm
  - a) en punktskattning av den sökta parametern
  - b) ett (approximativt) 95%-igt konfidensintervall för den sökta parametern. Vilka antaganden gör Du vid bestämningen av detta intervall?
2. Gör ett *stratifierat urval* av totalt 30 kommuner. Därvid skall Du först välja en lämplig stratifieringsvariabel (som alltså är en annan variabel än undersökningsvariabeln) och med hjälp av denna göra en indelning av samtliga 290 kommuner i 2-4 strata. Du väljer själv hur de olika strata skall definieras, dvs mellan vilka gränser *på stratifieringsvariabeln* som varje stratum av kommuner skall ligga. Därefter bestäms hur det totala urvalet av 30 kommuner skall fördelas mellan strata, och till sist görs ett *obundet slumpmässigt urval utan återläggning från varje stratum*. (Det kan då vara lämpligt att välja den ram som är sorterad efter den valda stratifieringsvariabeln.) Då erhålles de värden på undersökningsvariabeln som skall användas vid beräkningarna.
  - a) Motivera valet av de olika stegen vid stratifieringen samt ange den ram och de kommuner som urvalet kom att omfatta. Ange till sist också värdena på undersökningsvariabeln för dessa kommuner.
  - b) Bestäm en punktskattning av den sökta parametern.
  - c) Bestäm ett (approximativt) 95%-igt konfidensintervall för den sökta parametern. Vilka antaganden gör Du vid bestämningen av detta intervall?
  - d) Bestäm också längden på intervallet med stratifierat urval respektive intervallet med obundet slumpmässigt urval (uppgift 1) och jämför dessa. Vilket intervall är kortast? Varför?
3. Gör ett *systematiskt urval* av 30 kommuner.
  - a) Motivera valet av den ram Du har använt. Ange också de kommuner som Du har valt ut samt beskriv hur Du praktiskt har gått tillväga.
  - b) Bestäm en punktskattning av den sökta parametern.
  - c) Bedöm och motivera om Du tror att precisionen vid systematiskt urval har blivit lägre eller högre jämfört med ett obundet slumpmässigt urval av samma antal kommuner.
4. Vilket urvalsförfarande vill Du rekommendera i en praktisk situation? Motivera Ditt val!

De valda kommunerna, beräkningarna och resultaten redovisas i den separata svarsbilagan. Där skall också anges vilka startvärden på slumpalen som har använts vid de olika urvalen.

### Beskrivning av hur slumpalen används:

Med hjälp av Minitabs slumpalsgenerator finns på varje utdelat blad fem olika omgångar med 100 slumpal i vardera. Dessa är skapade *oberoende* av varandra med hjälp av olika starttal (BASE) och är placerade i kolumnerna C1-C5. Eftersom ramen med kommuner är numrerade, kan man koppla ihop ett sådant ordningsnummer med de slumpal som finns i respektive kolumn; dessa kommuner är då de som skall väljas i de olika urvalen. Läs slumpalen uppifrån och ner i respektive kolumn.

I denna uppgift skall tre olika urval göras, OSU, stratifierat urval med OSU inom varje stratum och systematiskt urval. Använd nu slumpalen i C1 till OSU-urvalet. Eftersom OSU-urvalen skall göras *utan återläggning*, gäller att om samma slumpal återkommer i kolumnen, hoppar man över detta tal och använder nästa.

I det stratifierade urvalet skall först två, tre eller fyra strata skapas från ramen och från varje stratum skall OSU-urval göras oberoende av varandra. Använd slumpalen i C2 för urvalet från stratum 1, C3 från stratum 2, C4 från stratum 3 och C5 från stratum 4 (vid behov). Om ett slumpal inte ligger i intervallet för aktuellt stratum hoppar ni över det talet. Detta utförs till dess att ni har uppnått vald urvalsstorlek inom respektive stratum.

Lämpliga slumpal till det systematiska urvalet kan tas från en godtycklig kolumn.