

**Obligatorisk examinationsuppgift SDA II, 3 högskolepoäng.  
Olika urvalsmetoder – punkt- och intervallskattningar**

Börja med att ladda ner "Kommuner2007.xls" från kursens hemsida.

Denna uppgift utföres individuellt. Sista dag för inlämning av uppgiften, vilket sker på en **speciell svarsbilaga** som laddas ner från kursens hemsida, är fredag 27 februari. Examinationen sker sedan vid seminarietillfälle S4. Svarsbilagan kan lämnas i därför avsedda lådfack på plan 7, B-huset, eller direkt till seminarieläraren. Glöm inte att skriva namn, gruppnummer och seminarielärares namn samt att ta en kopia på ditt inlämnade arbete!

*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 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. Nedan visas ett exempel på en sådan slumptalssida.

För var och en av kommunerna i populationen som återfinns i "Kommuner2007.xls" 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 skattekr) i kommunen 2007.

Materialet i "Kommuner2007.xls" 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

För att se de olika sätten att återge observationerna klickar ni på flikarna längst ner på excelarket.

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

### Uppgifter:

Antag att Du är intresserad av att veta vad den genomsnittliga kommunalskatten 2007 är. I praktiken skulle Du kunna undersöka kommunalskattesatserna för alla kommuner, men antag att Du istället skall göra en urvalsundersökning. I praktiken skulle Du då välja *ett* urvalsförfarande, men i denna uppgift kommer Du använda dig av tre urvalsförfaranden (OSU, stratifierat urval, samt systematiskt urval) för att få träning i dessa. Då Du genomför de tre urvalen skall Du besvara uppgifter 1, 2 och 3 på **svarsbilagan** som sedan skall lämnas in till seminarieläraren. I samtliga urvalsförfaranden är den sökta parametern ”Den genomsnittliga kommunalskatten 2007 i landets samtliga kommuner”. Nedan följer en beskrivning av vad som skall göras, men se också svarsbilagan.

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.) Sedan 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 en separat svarsbilaga. Där skall också anges vilka startvärden på slumptalen som har använts vid de olika urvalen.

### Beskrivning av hur slumptalen – se nästnästa sida – används:

Med hjälp av Minitabs slumptalsgenerator finns på varje utdelat blad – ett exempel finns på nästnästa sida - fem olika omgångar med 100 slumptal 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 slumptal som finns i

respektive kolumn; dessa kommuner är då de som skall väljas i de olika urvalen. Läs slumptalen 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 slumptalen i C1 till OSU-urvalet. (De första slumptalen är här 145, 7, osv men observera att Du har andra slumptal – se Din individuella slumptalstabell). Eftersom OSU-urvalen skall göras *utan* återläggning, gäller att om samma slumptal 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 slumptalen i C2 för urvalet från stratum 1, C3 från stratum 2, C4 från stratum 3 och C5 från stratum 4 (vid behov). Lämpliga slumptal till det systematiska urvalet kan tas från en godtycklig kolumn.

## **Betyg**

På denna uppgift ges något av betygen godkänd eller underkänd. Observera att för att kunna ges betyget godkänd måste uppgiften lämnas in senast fredag 27 februari. En uppgift kan behöva kompletteras för att uppnå betyget godkänd, denna komplettering måste färdigställas inom en veckas tid från seminarium 4. Är kompletteringen ej nöjaktigt utförd ges ytterligare en möjlighet, som ånyo måste utföras inom en veckas tid. Enbart två försök till komplettering ges alltså. För att överhuvudtaget behandlas vid första rättningen, efter inlämning 27 februari, måste allvarligt menade försök till lösning av uppgiftens samtliga huvuddelar ha gjorts, en första komplettering kan således ej bestå av att uppgiften s a s i princip görs från början. En undermålig uppgift kommer att direkt underkännas utan möjlighet till komplettering; uppgiften får då göras från början med en ny slumptalstabell.

```

MTB > base 1000
MTB > random 100 c1;
SUBC> integer 1 289.
MTB > base 1005
MTB > random 100 c2;
SUBC> integer 1 289.
MTB > base 1010
MTB > random 100 c3;
SUBC> integer 1 289.
MTB > base 1015
MTB > random 100 c4;
SUBC> integer 1 289.
MTB > base 1020
MTB > random 100 c5;
SUBC> integer 1 289.
MTB > print c1-c5

```

C1 används till OSU-urvalet  
C2 " " Stratum nr 1  
C3 " " Stratum nr 2  
C4 " " Stratum nr 3  
C5 " " Stratum nr 4 (vid behov)

Fortsättning på slumptalen

Rad	C1	C2	C3	C4	C5	Rad	C1	C2	C3	C4	C5
1	145	151	158	164	83	51	90	274	170	65	53
2	7	60	114	167	151	52	176	68	249	142	269
3	101	169	237	17	34	53	51	47	42	38	105
4	100	254	119	274	156	54	38	287	247	206	200
5	120	78	36	284	76	55	217	267	28	78	122
6	274	130	275	131	57	56	126	31	225	129	156
7	245	282	29	66	1	57	85	154	223	3	8
8	252	7	51	96	75	58	203	137	72	6	154
9	192	119	45	261	221	59	158	47	224	113	52
10	233	217	201	185	148	60	81	140	199	258	249
11	141	192	244	6	277	61	284	281	279	276	248
12	183	228	272	28	67	62	180	71	251	142	12
13	179	188	196	204	236	63	263	106	238	80	177
14	103	124	145	166	33	64	284	197	111	24	188
15	40	142	244	57	269	65	130	246	73	188	46
16	171	196	221	246	170	66	201	162	122	83	192
17	213	274	46	106	187	67	172	69	255	153	31
18	200	138	76	13	53	68	263	133	3	162	228
19	23	128	233	49	51	69	185	242	10	67	116
20	166	55	233	122	286	70	259	256	254	251	113
21	104	223	52	171	14	71	60	228	108	276	184
22	161	245	40	124	26	72	180	100	20	229	218
23	186	242	10	66	134	73	46	241	147	53	239
24	260	194	127	61	250	74	275	228	182	135	157
25	87	195	13	121	195	75	40	75	111	147	1
26	182	126	71	16	242	76	224	92	249	117	94
27	8	251	205	158	142	77	67	101	136	171	64
28	287	72	145	218	249	78	232	267	13	48	214
29	72	76	80	84	98	79	105	177	248	30	199
30	69	95	121	147	11	80	81	105	128	152	113
31	53	1	238	186	175	81	39	213	99	274	249
32	164	283	112	231	186	82	241	182	123	64	173
33	275	40	94	148	218	83	32	28	24	21	30
34	223	92	249	118	177	84	15	70	126	181	114
35	9	149	289	139	82	85	28	117	206	6	89
36	143	60	267	184	289	86	185	117	49	270	267
37	88	225	73	209	104	87	193	119	45	260	107
38	125	164	203	242	54	88	214	108	3	186	224
39	24	101	177	254	88	89	213	279	56	122	155
40	109	25	231	147	107	90	3	89	174	260	17
41	115	74	33	281	279	91	181	198	214	230	203
42	137	101	64	28	103	92	5	24	43	62	34
43	245	243	242	240	9	93	256	36	106	175	132
44	246	36	115	194	27	94	188	223	258	5	273
45	35	270	217	163	198	95	101	187	272	69	153
46	260	129	288	157	33	96	287	24	49	75	89
47	217	186	155	123	250	97	147	10	161	23	189
48	215	122	29	225	127	98	127	234	52	160	2
49	234	283	44	93	179	99	132	37	231	135	66
50	114	188	263	48	231	100	166	194	221	248	46

