

Tentamen

Linköpings Universitet, Institutionen för datavetenskap, Statistik

Kurskod och namn:	732G19 Utredningskunskap 1; 732G04 Surveymetodik
Datum och tid:	2013-10-19, 8 – 12
Jourhavande lärare:	Tommy Schyman
Tillåtna hjälpmedel:	Valfri räknedosa, till tentamen vidhäftad formelsamling.
Betygsgränser:	Tentamen omfattar totalt 20p. 12 poäng och uppåt ger betyget G, 16 poäng och uppåt ger betyget VG.

Redovisa och motivera tydligt alla dina lösningar och tolka alla dina svar!

Siffrorna i uppgifterna är påhittade.

Uppgift 1 (7p)

Antag att restaurangkedjan Subway har planer på att starta upp en restaurang i anslutning till Campus Valla. För att undersöka kundunderlaget vill man göra ett obundet slumpmässigt urval bland de 19 000 studenter som studerar på Campus Valla. Vi begränsar oss till frågan "Är det sannolikt att du kommer äta på Subway minst en gång i månaden?" som enbart har svarsalternativen *Ja* och *Nej*.

- Antag att Subway accepterar en felmarginal på 4 procentenheter för ett dubbelsidigt konfidensintervall (95 %) för andelen ja-sägare på den aktuella frågan. Hur stort stickprov bör man dra? (2p)
- Undersökningen genomförs med den beräknade stickprovsstorleken i a) och 56 % av de tillfrågade svarar *Ja* på frågan (dock är det ett bortfall på 30 %). Subway planerar att starta upp en restaurang om en majoritet av studenterna kommer äta där minst en gång i månaden. Undersök med hjälp av ett lämpligt konfidensintervall (99 % konfidensnivå) om Subway bör starta upp restaurangen. (2p)
- För att väga in bortfallets åsikt i frågan kontaktas 20 % av bortfallet på annat sätt än tidigare och man får svar från alla dessa. 47 % uppger att de skulle äta på restaurangen minst en gång i månaden. Ta hänsyn till denna (nya) information och skapa åter igen ett lämpligt konfidensintervall (99 % konfidensnivå) som besvarar frågan om Subway bör starta upp restaurangen eller ej. (3p)

Uppgift 2 (5p)

En högstadieskola vill undersöka hur mycket tid (i minuter) dess elever motionerar i veckan. I tabellen nedan finns antalet elever och uppskattningen på variansen i motionstid för de olika årskurserna.

Årskurs	Antal	Uppskattad σ_i^2
7	220	200
8	240	250
9	190	300

- a) Högstadieskolan accepterar en bredd på 8 minuter för ett dubbelsidigt konfidensintervall med konfidensnivån 95 %. Använd proportionell allokering för att bestämma hur många som ska undersökas i varje stratum. (2p)

Undersökningen genomförs med de beräknade urvalsstorlekarna och följande information sammanställs:

Årskurs	Genomsnittlig motionstid	Standardavvikelse
7	196	13,2
8	177	16,5
9	168	17,1

- b) Beräkna ett 95 % dubbelsidigt konfidensintervall för den genomsnittliga motionstiden bland högstadieskolans elever. (3p)

Uppgift 3 (6p)

Antag att högstadieskolan från uppgift 2 istället slumpmässigt valt ut fyra klasser av skolans totalt 22 och fått följande information:

Klass	Antal elever	Totalt antal motionsminuter
1	32	6320
2	27	4860
3	24	4700
4	31	6100

- a) Beräkna ett dubbelsidigt konfidensintervall (99 % konfidensnivå) för den genomsnittliga motionstiden bland högstadieskolans elever. (3p)
- b) Antag att urvalet gjorts med urvalssannolikheter proportionella mot klassernas storlekar. Beräkna åter igen ett dubbelsidigt konfidensintervall (99 %) för den genomsnittliga motionstiden. (3p)

Uppgift 4 (2p)

- a) Om man inte har någon kännedom om populationsandelen P vid urvalsdimensionering används $P = 0,5$. Varför? (1p)
- b) Ge ett exempel på en situation när PPS-urval av kluster kan vara att föredra framför OSU av kluster. (1p)