

Tentamen i Undersökningsmetodik (3.5 hp)

Kurs: Statistik och dataanalys II

2024-02-23

Skrivtid:	kl. 8.00 - 12.00 (4 timmar)
Godkända hjälpmedel:	Behövs inte men miniräknare utan lagrade formler och lagrad text är tillåtet
Vidhäftade hjälpmedel:	inga

- Tentamen består av tre uppgifter uppdelade i deluppgifter. Maximalt antal poäng anges per deluppgift.
- OBS! För var och en av deluppgifterna behöver du inte skriva långt och mycket, bara det viktigaste. Som vägledning kan du sikta på ca $\frac{1}{2}$ - 1 A4 per deluppgift. Inga beräkningar behövs, men du kan illustrera med bilder och skisser om det underlättar. Även formler, dvs. matematiska framställningar, går bra om du tycker att det är relevant och enklare.
- Maxpoäng är 100 och för godkänt resultat krävs minst 50 poäng. Betygsgränser:
 - A: 90 – 100
 - B: 80 – 89
 - C: 70 – 79
 - D: 60 – 69
 - E: 50 – 59
 - Fx: 40 – 49
 - F: 0 – 40
-
- Svar med fullständiga redovisningar och motiveringar ska lämnas.
 - Använd endast skrivpapper som tillhandahålls i skrivsalen.
 - Max en uppgift per blad, t.ex. uppgift 1a-d på ett eller flera blad men börja på nytt blad för nästa uppgift 2a-d osv.
 - För full poäng på en uppgift krävs tydliga, utförliga och väl motiverade förklaringar.
 - Kontrollera dina svar! Slarvfel kan ge poängavdrag!
- Lösningförslag läggs ut på Athena.

LYCKA TILL!

Uppgift 1. (30p)

- a) Under kursen har vi gått igenom fyra huvudtyper av s.k. *systematiska fel* som kan uppstå i en undersökning. Ange dessa och ge ett exempel för var och en av dessa hur de uppstår. Vilken huvudsaklig effekt på parameterskattningar medför dessa fel, en effekt som man bör vara särskilt orolig för? (10p)
- b) Förklara vad *tratt-tekniken* är. Risken för och effekten av två feltyper, se fråga a) ovan, kan antas minska om man använder tratt-tekniken. Vilka är de två feltyperna och varför påverkas just dessa? Och varför påverkas inte de två andra feltyperna? (10p)
- c) Förklara vad *imputering* är. När används det och varför? En variant kallas medelvärdesimputering, hur går det till. Och vilken effekt får medelvärdesimputeringar på variansskattningar och varför blir det så? (10p)

Uppgift 2. (40p)

- a) Under kursen beskrevs sju *kvalitetskriterier* som enligt lag ska beaktas och deklarerats (dvs. beskrivas) i undersökningar som ingår i Sveriges officiella statistik (SOS). Välj två av dessa kriterier och beskriv dem kortfattat. Förklara varför de är viktiga ur ett kvalitetsperspektiv, gärna med exempel. Beskriv också en målkonflikt som kan uppstå mellan de två kriterier som du valt. Om du inte tror att det finns en målkonflikt så måste det motiveras. (15p)
- b) Det har blivit allt vanligare med så kallade *webbpaneler*. Förklara vad en webbpanel är för något, hur man rekryterar/väljer individer till panelen och hur man typiskt använder information i panelen i en undersökning. Beskriv slutligen vad den huvudsakliga kritiken mot webbpaneler är. Argumentera för varför webbpaneler trots allt kan vara användbara. (15p)
- c) Svenska statistikfrämjandet har tagit fram en *etisk kod* för statistiker och statistisk verksamhet. Behövs det verkligen en sådan? Utveckla dina tankar lite kortfattat kring behovet av en etisk kod. Vilka är det som berörs av koden? Vad används statistiska undersökningar och analyser till och vilka är det som använder resultaten och i vilket syfte? (10p)

Uppgift 3. (30p) – Resonera utifrån ett case/kontext/sammanhang

Du ingår i en projektgrupp som har fått i uppdrag att studera användandet av AI-verktyg i undervisningen på gymnasieskolor i Sverige.

AI står för artificiell intelligens, och det finns ett antal olika verktyg t.ex. ChatGTP, Grammarly m.fl. Att använda AI som ett verktyg i undervisning är ett relativt nytt fenomen det finns inte så mycket tidigare forskning kring detta.

Ni har fått ett halvår på er innan undersökningen ska genomföras och innan dess måste ni reda ut ett antal olika frågor kring undersökningsdesignen, t.ex. operationalisera ett antal frågeställningar och utifrån detta konstruera ett helt nytt frågeformulär.

- a) Hur skulle du karaktärisera den här studien? Är det en deskriptiv, förklarande, prediktiv eller preskriptiv studie? Är det en tvärsnitts- eller tidsseriestudie? Experimentell studie eller observationsstudie av surveytyp? Förklara kortfattat. (6p)
- b) Hur ska målpopulationen definieras? Vad skulle du använda för ram? Vilka ska urvalsubjekten vara och vilka är undersökningsobjekten? (10p)
NOT: När du ger förslag på en lämplig ram så försök vara realistisk. Notera att varken RTB (Registret över totalbefolkningen) eller SPAR (Statens personadressregister) innehåller uppgifter om en viss person går på gymnasiet eller om man lärare. Hur ska ni göra för att nå de ni vill undersöka?
- c) Vilken datainsamlingsmetod skulle du förslå för den här studien? Ange minst två skäl till varför du väljer den metoden. (4p)
- d) I ett första utkast till frågeformuläret hittar du följande fråga med tillhörande svarsalternativ som en kollega har lagt till:

Jag har erfarenhet av AI och anser att det utgör ett konstruktivt komplement till den lärarledda undervisningen.

- Ja Nej Kanske

Vilka problem kan du identifiera med den här frågeformuleringen och svarsalternativen? Föreslå en ny version av frågan och svarsalternativen. Gör om och gör rätt så att säga! (10p)