

Uppgift 1(12p)

Kontroll av nykterheten hos bilförare på svenska vägar görs som du vet av polisen. Man stoppar ett fordon och föraren får blåsa i en s.k. alkometer. Ger apparaten utslag för alkohol, får föraren lämna ett blodprov hos läkare.

Blodprovet ger juridiskt bindande bevis för rattfylleri medan resultatet i det test polisen gör endast uppfattas som en indikation på rattfylleri. Det finns alltså en osäkerhet i polisens test och fel kan begås på två sätt:

- i) en nykter person kan klassificeras som onykter
- ii) en onykter person kan klassificeras som nykter.

Låt oss anta att, vid en speciell tidpunkt under veckan, 5 % av alla bilförare som kör på vägarna har mer alkohol i blodet än vad som är tillåtet. Anta vidare att den mätmetod som polisen använder i 9 fall av 10 ger rätt resultat, dvs både i att visa att en rattonykter person är rattonykter och att en nykter person är nykter.

När en polis stoppar ett fordon för nykterhetskontroll kan föraren vara antingen nykter eller ej nykter och i testet kan personen klassificeras som antingen nykter eller ej nykter.

- a) Bestäm i detta fall prevalens, sensitivitet och specificitet.
- b) Bestäm sannolikheten att en person som av polisen bedöms som onykter faktiskt är nykter.

Uppgift 2 (7p)

Kommentera nedanstående påståenden med antingen **Sant** eller **Falskt**. Du behöver ej motivera ditt svar.

- a) Medianen kan aldrig vara större än tredje kvartilen.
- b) Sannolikheten att få två sexor direkt efter varandra vid två upprepade kast med en tärning är $1/36$.
- c) Genom att öka konfidensgraden blir konfidensintervallet kortare.
- d) Ett stapeldiagram är lämpligt för att illustrera fördelningen över en kvalitativ variabel.
- e) Pearsons korrelationskoefficient är lämplig att använda då variablerna mäts på ordinalnivå.
- f) Samplingfördelningen för urvalsmedelvärdet är approximativt normalfördelad under förutsättning att urvalet är slumpmässigt draget, ursprungsfördelningen ej är extremt sned och urvalsstorleken är större än 30.
- g) Temperatur i grader Celcius mäts på ordinalnivå.

Uppgift 3 (10p)

För att skatta andelen individer med en viss sjukdom i en stor population ska ett slumpmässigt urval om ett antal individer dras så att längden av ett 95 % konfidensintervall inte får överstiga 0.06.

- Hur stort stickprov måste dras? Utgå från att den sanna andelen inte överstiger 0.2.
- För vilka värden på andelsskattningen kommer det beräknade intervallet att bli kortare än 0.06? Motivera kort!

Uppgift 4 (12p)

I en studie vill man bland annat studera om kolesterolvärde (mg/100 ml plasma) kan förklaras av variablerna vikt (kg) och ålder (år). Man skattar därför en regressionsmodell med kolesterolvärde som beroende variabel och vikt och ålder som förklarande variabler. Resultatet presenteras nedan.

- Vilken av de två förklaringsvariablerna anser du vara den viktigaste prediktorn för kolesterolvärde? Motivera!
- Ange och tolka förklaringsgraden.
- Prediktera kolesterolvärdet för en 40-årig patient som väger 70 kg med hjälp av den skattade modellen.
- Tolka, i de fall som är meningsfulla, de skattade parametrarna i modellen.

Regression Analysis

The regression equation is

$$\text{Kolesterol} = 67.5 + 0.558 \text{ Vikt} + 5.23 \text{ Ålder}$$

Predictor	Coef	StDev	T	P
Constant	67.52	54.38	1.24	0.227
Vikt	0.5583	0.7233	0.77	0.448
Ålder	5.2264	0.7409	7.05	0.000

S = 43.85

R-Sq = 70.9%

R-Sq(adj) = 68.3%

Uppgift 5 (9p)

Begreppet p-värde är centralt inom statistisk teori.

- a) Hur definierar du allmänt ett p-värde?
- b) Vilket eller vilka av följande påståenden beträffande ett p-värde lika med 0.002 är korrekt? Ingen motivering behövs.
 - 1) Det låga p-värdet indikerar att nollhypotesen är sann.
 - 2) Sannolikheten att nollhypotesen är sann är mindre än 0.002.
 - 3) Vi har mycket starkt stöd från datamaterialet för att alternativhypotesen är sann eftersom p-värdet är så lågt.
- c) Kan man säga att p-värdet är den lägsta signifikansnivån nollhypotesen kan förkastas på? Motivera!