

Tentamen i Statistisk analys

14 januari 2026 kl. 8-13

Examinator: Tom Britton, tel. 08-16 45 34, tom.britton@math.su.se

Tillåtna hjälpmedel: Formel- och tabellsamling (delas ut) och miniräknare.

Återlämning: Tentan kommer vara rättad senast två veckor efter tentamensdagen. När tentan är rättad meddelas detta via kursforumet och därefter kan tentan hämtas ut vid matematik-expeditionen.

Varje korrekt löst uppgift ger 10 poäng. Gränsen för godkänt är preliminärt 30 poäng. För D krävs 34 p, för C 40 p, för B 48 p och för A krävs 54 p. I dessa poäng ingår eventuella bonuspoäng. Texten ska vara väl läsbar och resonemang ska vara klara och tydliga.

Det skall tydligt framgå hur beräkningar gjorts. Kommunikation med andra personer är ej tillåtet och kommer anmälas vid uppdragande.

Uppgift 1

Nedan följer 5 påståenden att svara sant eller falskt på (eller ingenting om man inte vet). Korrekt svar på respektive påstående ger 2p, fel svar ger -2p och inget svar ger 0p (om totalsumman skulle bli negativ sätts poängen till 0).

- a) Icke-parametriska metoder kan användas även om data är normalfördelat men har då lite lägre styrka än parametriska metoder.
- b) En fördel med observationsstudier jämfört med randomiserade studier är att man slipper slumpa individer till olika grupper.
- c) Styrkan för ett hypotestest (för en given situation) ökar ju fler observationer som testet baseras på.
- d) Lådagram (Eng: Box plot) är lämpliga när flera olika datamaterial ska jämföras.
- e) t -fördelningen har tjockare svansar ju fler frihetsgrader den har.

Uppgift 2

Vid en undersökning av två näringsextrakt uppmäts tid till att blommor vissnade. Det första extraktet testades på sex blommor och det andra extraktet på sex andra blommor (vilken

blomma som fick vilken näring avgjordes med lottning). Resultatet, med dygn som tidsenhet, blev för blommorna med första extraktet: $x_1 = 3.1, x_2 = 2.2, x_3 = 4.5, x_4 = 3.9, x_5 = 4.9, x_6 = 3.6$, och med andra extraktet: $y_1 = 5.1, y_2 = 3.2, y_3 = 4.0, y_4 = 4.4, y_5 = 5.4, y_6 = 3.7$. Följande gäller för angivna data: $\bar{x} = 3.7, \bar{y} = 4.3, \sum_i x_i^2 = 86.88, \sum_i y_i^2 = 114.46$. Antag att observationerna kan betraktas som någorlunda normalfördelade och med samma standardavvikelse σ .

a) Testa hypotesen att de två extrakten har samma väntevärde på 95%-nivån, under förutsättning att $\sigma = 1$. (4 p)

b) Gör motsvarande test där σ skattas. (6 p)

Uppgift 3

Temperaturen på julskinkan i ugnen kan antas öka hyfsat linjärt under dess ugnstekning. Temperaturen uppmättes, genom stick in i skinkan på olika platser i mitten, varje halvtimme sedan den sattes in och gav följande temperaturer: 0.5h: 17, 1h: 26, 1.5h: 35, 2h: 47, 2.5h: 58. Vi antar att variationen beror på stickens placering och att det inte finns något beroende mellan mätvärden nära i tid.

a) Hur mycket ökar temperaturen per timme, och vilken *förväntad* temperatur hade julskinkan när den sattes in i ugnen (ange skattning och 95% konfidensintervall). (7 p)

b) Prediktera vilken observerad temperatur skinkan kommer ha vid 3h, dvs ange ett 95% prediktionsintervall (under förutsättning att extrapolation fungerar) (3 p)

Uppgift 4

Vid en doktorandkonferens med deltagare från Sverige och Norge fick doktoranderna göra ett diagnostiskt test som därefter graderas med något av betygen U, G eller VG. Utfallet blev enligt följande tabell:

Land Betyg	U	G	VG	Totalt
Sverige	7	16	11	34
Norge	12	14	8	34
Totalt	19	30	19	68

Testa om det föreligger någon signifikant skillnad mellan doktorandernas kunskaper i de två länderna. (10 p)

Uppgift 5

Ett stickprov $x_1, \dots, x_{10}$ från exponentialfördelning med parameter β samlas in. Antag att medelvärdet av observationerna blir $\bar{x} = 3.2$.

a) Härled med hjälp av likelihooden vad ML-skattningen av β blir. (4 p)

b) Vad blir ML-skattning för väntevärdet $E(X)$ (motivera ditt svar)? (2 p)

- c) Vilken varians har skattningen från b). Använd detta uttryck för att konstruera ett approximativt 95% konfidensintervall för skattningen i b). (4 p)

Uppgift 6

För att undersöka om Svenska storföretags vd:ars ålder påverkar vilken lön de har erhöles följande ålder och årslöner för 10 stora svenska bolag (ålder i år och årslön i MSec), sorterade i åldersordning: (35, 8.8), (40, 7.1), (44, 15.0), (47, 13.2), (52, 16.9), (56, 15.7), (59, 18.0), (62, 17.2), (65, 40.4), (69, 25.0). (Påhittade data.)

- a) Varför är Pearsons rangkorrelation olämplig att använda i detta fall? (2 p)
- b) Beräkna Spearmans rangkorrelation och testa på 99%-nivån om det finns ett statistiskt säkerställt samband mellan lönen och vd:ars ålder. (8 p)

Lycka till!