Matematisk statistik	**Tentamen: 2016–08–17 kl 14^{00}–19^{00}**
Matematikcentrum	**FMS 086 — Matematisk statistik för B, K, N och BME, 7.5 hp**
Lunds tekniska högskola	**MASB02 — Matematisk statistik för kemister, 7.5 hp**
Lunds universitet	

Korrekt och väl motiverad lösning på uppgifterna 1–5 ger 10 poäng vardera medan delfrågorna på uppgift 6 ger 4 poäng vardera. Totalt kan man få 70 poäng. Gränsen för godkänd är 35 poäng, dock finns det vissa minimikrav på uppgifterna 1–5 (18p) respektive uppgift 6 (7p).

Institutionens papper används både som kladdpapper och inskrivningspapper. Varje lösning skall börja överst på nytt papper. Rödpenna får ej användas.

Tillåtna hjälpmedel: Matematiska och statistiska tabeller som ej innehåller statistiska formler, formelsamling matematisk statistik för bio- och kemitekniker, samt miniräknare.

Resultatet förs in i LADOK senast fredag den 2 september i matematikhusets entréhall.

Redovisa införda beteckningar; ange modeller, approximationer, hypoteser och slutsatser. Motivera alla antagande.

Skriv anonymkoden och personlig identifierare (eller namn) på omslaget och SAMTLIGA inlämnade papper.

---

1. I ett visst system ingår en komponent som har en exponentialfördelad livslängd med väntevärde 30 timmar. När komponenten går sönder skickas en felsignal till en kontrollör. Tiden från felsignal till dess komponenten är utbytt kan beskrivas med en slumpvariabel som är rektangelfördelad i intervallet (1, 2) timmar.

   (a) Bestäm väntevärde och standardavvikelse för den tid, T, en komponent är i systemet. (5 p)

   (b) Man har 25 komponenter i lager. Beräkna sannolikheten att dessa komponenter räcker i mer än 5 veckor, d.v.s. $5 \cdot 7 \cdot 24 = 840$ timmar. (5 p)

2. Två olika metoder, för att mäta små halter av nonylfenoletoxylat i bomullskläder, jämfördes. Prover från 6 olika tröjor togs, varje prov delades upp i två delar och analyserades med respektive metod. Följande resultat erhölls:

Prov nr	1	2	3	4	5	6
Metod A	0.18	0.09	0.21	0.17	0.30	0.19
Metod B	0.22	0.12	0.22	0.16	0.34	0.22

   Antag lämlig(a) normalfördelning(ar) och testa på nivån 1 % om det finns någon systematisk skillnad mellan metoderna. (10 p)

3. Enligt EUs riktlinjer för luftföroreningar så bör dygnsvärdet av marknära ozon inte överskrida 120 μg/m^3 mer än 25 gånger per år. Under två veckor i juni (14 dagar) överskreds gränsvärdet 6 gånger.

   (a) Undersök, med **antagande om oberoende mellan dagarna**, om detta antyder att överskridande sker oftare än 25 dagar per år. (7 p)

   (b) Eftersom marknära ozon är kopplat till höga temperaturer sker nästan alla överskridande under sommaren. Pelle tycker därför att man borde använda 25 överskridande under 92 dagar (juni, juli, augusti) som referens. Påverkar detta slutsatsen? (3 p)

4. I en grupp om 32 kvinnor med förhöjd kolesterolhalt ville man pröva en ny kolesterolsänkande medicin (medicin B). Man valde slumpmässigt ut 16 av kvinnorna som fick den nya medicinen. De övriga fick ett annat, väl beprövat, läkemedel (medicin A). Efter en tid av medicinering mätte man förändringen, y_i, (före-efter) hos samtliga 32 kvinnor. Eftersom kvinnornas ålder förmodligen påverkar förändringen i halt ansatte man följande regression modell

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \qquad \varepsilon_i \in N(0, \sigma^2) \text{ oberoende}$$

där x_{1i} är 0 för medicin A och 1 för medicin B och x_{2i} är kvinnornas ålder. Följande värden har beräknats för datamaterialet

$$\beta^* = \begin{pmatrix} 2.418 \\ 2.114 \\ -0.048 \end{pmatrix} \qquad s^2 = 0.253 \qquad \left(X^T X\right)^{-1} = \begin{pmatrix} 1.068 & -0.073 & -0.019 \\ -0.073 & 0.125 & 2.087 \cdot 10^{-4} \\ -0.019 & 2.087 \cdot 10^{-4} & 3.711 \cdot 10^{-4} \end{pmatrix}$$

(a) Hur tolkas skattningarna av de tre parametrarna β_0, β_1, och β_2? (2 p)

(b) Avgör, på 5%-nivån, om det finns en ålderseffekt. (5 p)

(c) Avgör, på 5%-nivån, om den nya medicinen är **bättre** (dvs ger en större förändringen av kolestrolhalten) än medicin A. (3 p)

5. För att undersöka hur slitytan hos klinkerplattor påverkas av en ny färgtillsats tänker Emma tillverka 10 plattor med det nya färgtillsatsen och 10 plattor med en gammal beprövad färgtillsats. Därefter slipas alla 20 plattorna i 1 minut under kontrollerade former och mängden avnött material (mm^3) mäts. Tyvärr ryms endast 10 plattor i laboratoriets ugn. Emma väljer slumpmäsigt ut 5 plattor behandlade med den nya tillsatsen och 5 behandlade med den gamla till första bränningen, de resterande $5+5$ plattorna bränns i andra omgången. En rimlig modell för data är

$$y_{N1k} \in \mathsf{N}(\mu_1 + \Delta, \sigma^2) \qquad y_{G1k} \in \mathsf{N}(\mu_1, \sigma^2) \qquad y_{N2k} \in \mathsf{N}(\mu_2 + \Delta, \sigma^2) \qquad y_{G2k} \in \mathsf{N}(\mu_2, \sigma^2)$$

där index N, G anger vilken färgtillsats som använts, index 1, 2 anger vilken omgång plattan bränts under och $k = 1, \ldots, 5$ numrerar plattorna. En skattning av Δ är

$$\Delta^* = \frac{\bar{y}_{N1} + \bar{y}_{N2} - \bar{y}_{G1} - \bar{y}_{G2}}{2}$$

Data från experiment är

$$\begin{array}{llll} \bar{y}_{N1} = 15.77 & \bar{y}_{N2} = 13.93 & \bar{y}_{G1} = 11.84 & \bar{y}_{G2} = 10.53 \\ s^2_{N1} = 8.84 & s^2_{N2} = 9.63 & s^2_{G1} = 17.05 & s^2_{G2} = 9.85 \end{array}$$

(a) Vad kallas den försöksplan som Emma använder? (2 p)

(b) Vad är väntevärdet och variansen för Δ^* (uttryckt i μ_i, Δ och σ)? (2 p)

(c) Skatta σ^2. (2 p)

(d) Avgör om det finns någon signifikant skillnad mellan tillsatserna. (4 p)

6. Teorifrågor. Ge koncisa svar på nedanstående frågor.

(a) Vid en laboration i kemi ska studenterna utföra ett tekniskt svårt experiment. Baserat på erfarenhet från tidigare år lyckas experimentet i 70% av fallen. Om studenterna får göra om experimentet tills de lyckats en gång, vad är sannolikheten att det krävs högst 3 experiment? (4 p)

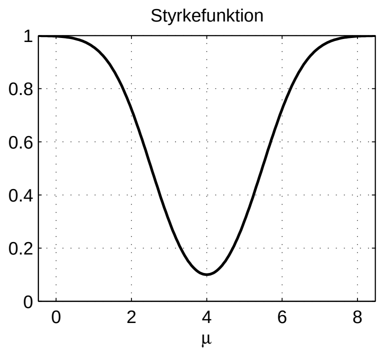
(b) En biolog studerar hur en viss typ av maskar fördelar sig i sand. Hon delar in en bit strand i rutor och räknar antalet maskar per ruta. Baserat på sina data uppskattar hon följande sannolikhetsfunktion för X, antal maskar per ruta.

Antal maskar per ruta	0	1	2	3	4	5
Sannolikhet	0.4	0	0	0.2	0.3	0.1

Beräkna väntevärde och varians för X. (4 p)

(c) För att testa $H_0 : \mu = 3$ mot $H_1 : \mu \geq 3$, på nivå 5 % i en $N(\mu, 2^2)$ förkastas H_0 om $\bar{x} > 3.85$. Hur många mätningar användes? (4 p)

(d) Styrkefunktionen för ett hypotestest baserat på n observationer från en normalfördelning med känd varians visas i figuren nedan.



I) Vilka hypoteser (H_0 och H_1) testas? (2 p)

II) Illustrera, med en **förenklad** figur, hur styrkefunktionen ändras om man **ökar** antalet observationer. (2 p)

(e) För att undersöka utbytet från en kemisk reaktion genomförs mätningar utan eller med katalysator (effekt A) och vid låg eller hög temperatur (effekt B). Experimentet kan beskrivas som ett 2^2-försök med 3 mätningar av varje faktorkombination och lämpliga antagande om normalfördelade fel.

I) Vilken kvantil används för att testa om en effekt är signifikant på 5%-nivån? (2 p)

Antag att samspelet inte är signifikant, vilka slutsatser kan man då dra från experimentet:

II) Om huvudeffekt A är signifikant? (1 p)

III) Om huvudeffekt B är signifikant positiv? (1 p)

LYCKA TILL!