

Skrivtid: 08.00-13.00. För betygen 3, 4 resp. 5 krävs 18, 25 resp. 32 poäng, inklusive ev. bonuspoäng. Lösningarna skall vara väl motiverade. Skriv endast på ena sidan, börja ny uppgift på ny sida och använd ej rödpenna.

Tillåtna hjälpmedel: Räknedosa. Formel- och tabellsamling i matematisk statistik.

1. Följande är ett oberoende stickprov från $N(\mu, \sigma^2)$, där både μ och σ är okända:

0.63, 0.36, -0.45, -0.017, 0.37, 1.99, 1.58.

- (a) Beräkna ett 99% konfidensintervall för μ . (4p)
(b) Man har en misstanke att $\mu = 0.2$. Använd konfidensintervallet för att uttala dig om detta med statistisk skärpa. (1p)
2. Man har funnit att antalet driftstopp X under en månad i en produktionsavdelning har följande sannolikhetsfördelning:

x	0	1	2	3
$p_X(x)$	0.55	0.25	0.15	0.05

- (a) Beräkna sannolikheten för minst ett driftstopp under en månad. (1p)
(b) Beräkna sannolikheten för exakt två driftstopp under en månad, givet att minst ett driftstopp inträffat den aktuella månaden. (1p)
(c) Produktionen innebär vissa kostnader, vilka delvis beror av antalet driftstopp. I fallen med 0 resp. 1 driftstopp är kostnaden i vardera fallet 10 000 kr; för fallen med 2 resp. 3 driftstopp är kostnaden 25 000 kr resp. 40 000 kr. Ange förväntad kostnad under en månads drift. (3p)
3. Låt $X_1, \dots, X_5$ samt $Y_1, \dots, Y_4$ vara oberoende slumpvariabler, sådana att alla $X_i \sim N(1, 2)$ och alla $Y_i \sim N(-2, 1)$. Beräkna följande sannolikhet: (5p)

$$P(X_1 + \dots + X_5 > Y_1 + \dots + Y_4).$$

4. (a) I figur 1 (sista sidan) visas lådagram för fyra stora stickprov, simulerade från de kontinuerliga fördelningarna (i)–(iv) nedan.
Matcha rätt lådagram (A,B,C,D) till rätt täthetsfunktion ((i),(ii),(iii),(iv)): (3p)

$$\begin{aligned} \text{(i)} \quad f(x) &= \frac{1}{1.5\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-10)^2/2.25\sigma^2}, -\infty < x < \infty, & \text{(ii)} \quad f(x) &= 0.05, 0 < x < 20, \\ \text{(iii)} \quad f(x) &= \frac{1}{10} e^{-x/10}, x > 0, & \text{(iv)} \quad f(x) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-10)^2/\sigma^2}, -\infty < x < \infty. \end{aligned}$$

- (b) För data i lådagram B, gör en grov skattning av standardavvikelsen (*Ledning*.
Tänk på kvantiler hos normalfördelningen.) (2p)
5. Energiförbrukningen per dygn under vintermånaderna har, för en viss villa med direkt eluppvärmning, väntevärde 68 kWh och standardavvikelse 12.5 kWh. Man är intresserad av att studera villans sammanlagda energiförbrukning under vintermånaderna december, januari och februari.

- (a) Är det rimligt att antaga oberoende i energiförbrukning mellan olika vinterdagar?
Motivera! (1p)
- (b) Gör antagande om oberoende och beräkna sannolikheten att villans sammanlagda energiförbrukning under december, januari och februari inte överstiger 6000 kWh.
(4p)
6. Vid ett försök mättes tryckhållfastheten hos 30 betongkuber gjutna med samma betongblandning. Mätresultaten i MPa:

26.00	27.26	27.55	27.60	28.65	28.88	29.68	29.82	30.80	30.88
31.00	31.20	31.20	31.50	32.25	33.51	34.45	35.01	35.50	35.60
35.80	35.86	36.52	36.60	37.10	37.80	38.00	38.00	38.50	39.01

Blandningen påstods ha 30 MPa som karakteristisk tryckhållfasthet, vilket innebär att som mest 5 % av alla betongkuber bör ha en tryckhållfasthet som understiger 30 MPa. Beräkna ett konfidensintervall för andelen kuber vars tryckhållfasthet understiger 30 MPa. Använd intervallet för att uttala dig om huruvida betongblandningens karakteristiska tryckhållfasthet verkligen är 30 MPa. (5p)

7. I en studie undersöktes kvicksilverhalt i 120 sjöar i Maine, USA. För varje sjö registrerades kvicksilverhalt och höjd över havet (elevation).

Samtliga observationer finns uppritade i figur 2 (sista sidan), och ett utdrag av data redovisas nedan.

<i>Sjö</i>	<i>Halt Hg (ppm)</i>	<i>Elevation (ft)</i>
Allen Pond	1.080	425
Alligator Pond	0.025	1494
Anasagunticook Lake	0.570	402
Balch & Stump Ponds	0.770	557
⋮	⋮	⋮
Wood (Little Big) Pond	0.250	1244

- (a) Skatta parametrarna α och β i modellen

$$Y_i = \alpha + \beta x_i + \epsilon_i, \quad i = 1, \dots, n$$

där Y_i betecknar kvicksilverhalt och x_i höjd över havet för den i :e sjön, och vidare $\epsilon_i \sim N(0, \sigma^2)$ antages vara oberoende slumpvariabler. (2p)

Räknehjälp: Använd följande sammanställning från samtliga 120 sjöar:

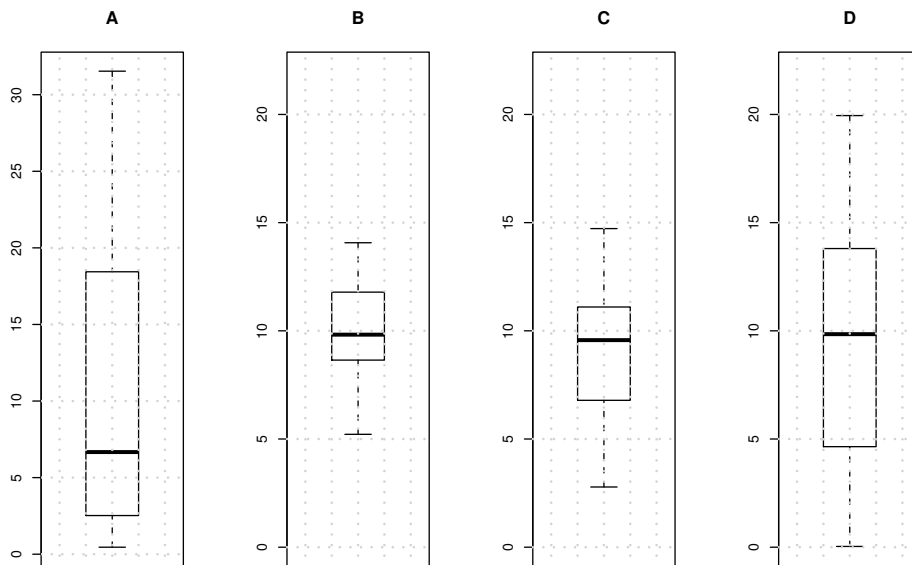
$$\sum_{i=1}^{120} x_i = 68074, \quad \sum_{i=1}^{120} y_i = 58.206,$$

$$\sum_{i=1}^{120} (x_i - \bar{x})(y_i - \bar{y}) = -5630.652, \quad \sum_{i=1}^{120} (x_i - \bar{x})^2 = 22073176, \quad \sum_{i=1}^{120} (y_i - \bar{y})^2 = 13.21695.$$

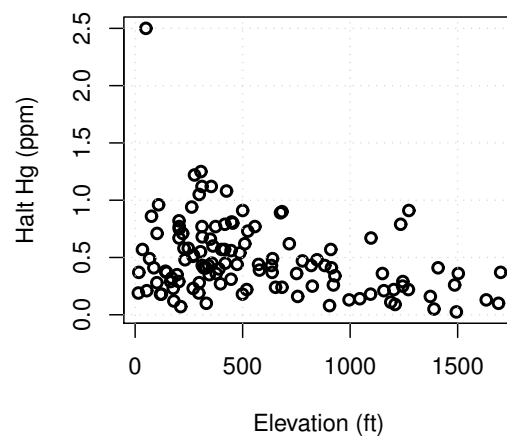
- (b) En viss sjö, ej med i ursprungsmaterialet, är belägen 500 ft över havet. Ange förväntad kvicksilverhalt. (1p)
- (c) Beräkna ett 95%-konfidensintervall för β . Använd detta för att motivera eventuell inverkan av elevation på kvicksilverhalten. (2p)

8. Vi har 10 oberoende, symmetriska slantar, och utför följande. Först singlar vi alla slantar en gång; sedan singlar vi på nytt alla de som visade krona efter första singlingen. Beräkna väntevärdet för det totala antalet klave man ser. (5p)

Lycka till!



Figur 1: Lådagran (problem 4).



Figur 2: Kvicksilverhalt i sjöar i Maine (problem 8).