Korrekt, väl motiverad lösning på uppgift 1–3 ger 10 poäng vardera medan uppgift 4–6 ger 20 poäng vardera. Totalt kan man få 90 poäng. Gränsen för godkänd är 40 poäng.

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

Tillåtna hjälpmedel: Matematiska och statistiska tabeller som ej innehåller statistiska formler, Formelsamling i matematisk statistik AK 1996 eller senare, samt miniräknare.

Resultatet anslås senast måndagen den 30 april i matematikhusets entréhall.

1. (a) Antag att $\mathsf{P}(A) = 0.2$, $\mathsf{P}(B) = 0.4$ och $\mathsf{P}(A \cup B) = 0.52$. Är händelserna A och B oberoende? (2p)

   (b) För en viss typ av malm vill man bestämma halten av ett givet ämne. Man tar därför ett prov av malmen som innehåller halten X av ämnet. Notera att X är en stokastisk variabel eftersom olika prover har olika sammansättning. Provet delas i två lika stora delar, och mätningar Z_1 och Z_2 av den sökta halten utförs på respektive del. Vi antar att $Z_i = X + Y_i$, där X är den verkliga halten och Y_i är mätfelet i den i:te mätningen. Vidare antar vi att X, Y_1 och Y_2 är oberoende stokastiska variabler med $\mathsf{E}(X) = 0.2$, $\mathsf{D}(X) = 0.04$, $\mathsf{E}(Y_i) = 0$ och $\mathsf{D}(Y_i) = 0.03$. Bestäm korrelationen $\rho(Z_1, Z_2)$. (4p)

   (c) Låt X vara rektangelfördelad mellan 0 och 1. Vilken fördelning får $Y = -4 \ln X$? (4p)

2. Vid tillverkning av klinkers blir dessa defekta, obeorende av varandra, med sannolikheten 0.01. Man garanterar att ett parti om 10000 klinkers innehåller högst 120 felaktiga. Bestäm approximativt sannolikheten för att det inte finns fler defekta än vad som anges i garantin. (10p)

3. En signal som antar värdet 0 med sannolikheten $\frac{2}{3}$ och 1 med sannolikheten $\frac{1}{3}$ sänds ut. Den passerar därefter två stationer som var och en återger inkommande signal felaktigt med sannolikheten $\frac{1}{7}$. Stationerna gör rätt eller fel oberoende av varandra. Vid mottagandet får man signalen 0. Bestäm den betingade sannolikheten för att det verkligen var signalen 0 som sänts. (10p)

4. Vid tillverkning av en produkt är man intresserad av att utvärdera skillnaden mellan två olika tillverkningsmetoder. Antag att vi har oberoende observationer $x_1, x_2, \ldots, x_n$ av $X \in N(\mu_x, 10)$ och $y_1, y_2, \ldots, y_n$ av $Y \in N(\mu_y, 10)$. Vid en preliminär undersökning har vi med $n = 20$ fått $\bar{x} = 23.2$ och $\bar{y} = 21.1$.

   (a) Testa

   $$\begin{aligned} H_0: &\quad \mu_x - \mu_y = 0, \\ H_1: &\quad \mu_x - \mu_y > 0. \end{aligned}$$

   på nivån 0.05. (10p)

   (b) Låt $\mu = \mu_x - \mu_y$. Ange styrkefunktionen $h(\mu)$ för testet i a). Hur stort måste n vara för att styrkefunktionens värde i punkten $\mu = 2$ skall vara större än 0.5? (10p)

5. Låt y' vara produktionen av någon gröda och x' mängden tillfört kväve mätt i någon lämplig enhet. Vi har följande mätningar:

x'	0.09	0.32	0.69	1.51	2.29	3.06	3.39	3.63	3.77
y'	16.4	66.7	109.9	181.3	191.5	193.2	181.3	127.7	200.3

(Observera att antalet observationer egentligen är i minsta laget för de nedanstående analyserna.)

Vi tror på följande samband mellan x' och y',

$$y' = e^{\alpha'} e^{\beta_1 x'} (x')^{\beta_2}.$$

Därför arbetar vi med den linjära modellen,

$$\ln Y_i' = \alpha' + \beta_1 x_i' + \beta_2 \ln x_i' + \varepsilon_i, \tag{1}$$

där $\varepsilon_1, \ldots, \varepsilon_9$ är oberoende och $N(0, \sigma)$. Om vi som vanligt betecknar de förklarande variablerna med $x_{i1} = x_i'$ och $x_{i2} = \ln x_i'$ och låter responsvariabeln vara $Y_i = \ln Y_i'$ så kan vi analysera datamaterialet enligt modellen,

$$Y_i = \alpha + \beta_1(x_{i1} - \bar{x}_{.1}) + \beta_2(x_{i2} - \bar{x}_{.2}) + \varepsilon_i\,; \; i = 1, \ldots, 9, \tag{2}$$

där $\bar{x}_{.1} = 2.083$ och $\bar{x}_{.2} = 0.253$. Nedanstående beräkningar gäller för den här modellen:

$$(X^T X)^{-1} = \begin{pmatrix} 0.1111 & 0 & 0 \\ 0 & 0.3858 & -0.3999 \\ 0 & -0.3999 & 0.4877 \end{pmatrix},$$

$\alpha^* = 4.7517, \beta_1^* = -0.4902, \beta_2^* = 1.0879$ och $Q_0 = 0.13467$.

(a) Bestäm konfidensintervall för β_1 i (1) med konfidensgraden 95%. (4p)

(b) Bestäm konfidensintervall för α' i (1). Låt konfidensgraden även här vara 95%. Ledning: Uttryck först $(\alpha')^*$ med hjälp av α^*, β_1^*, β_2^*, $\bar{x}_{.1}$ och $\bar{x}_{.2}$. (6p)

(c) En person påstår att om man tillför mängden en enhet kväve ($x' = 1$), så kommer väntevärdet för Y' vara minst 120. Stämmer detta med vårt regressionsmodellantagande? Avgör frågan genom att konstruera ett lämpligt intervall med hjälp av regressionsanalysen för modellen (2). (10p)

6. Låt $x_1, x_2, \ldots, x_n$ vara oberoende observationer från en stokastisk variabel X med sannolikhetsfunktionen,

$$p_X(k) = (\frac{\vartheta}{2})^{|k|}(1 - \vartheta)^{1-|k|}, \; k = -1, 0, 1; \; 0 \leq \vartheta \leq 1.$$

(a) Bestäm ML-skattningen av ϑ. (8p)

(b) Beräkna $\mathsf{E}(|X|)$ och $\mathsf{V}(|X|)$. (4p)

(c) Visa att ML-skattningen är väntevärdesriktig. (4p)

(d) Bestäm variansen för punktskattningen i (a). (4p)

Lycka till!