Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF1917/SF1919 SANNOLIKHETSTEORI OCH STATISTIK,
ONSDAG 10 JANUARI 2024 KL 8.00–13.00.

Examinator: Mykola Shykula, 08-790 6644.

Tillåtna hjälpmedel: Formel- och tabellsamling i Matematisk statistik (utdelas vid tentamen), miniräknare.

Tentamen består av två delar, benämnda del I och del II.
Del I består av uppgifterna 1–12 och varje korrekt svar ger 1 poäng. På denna del ska endast svar anges, i form av val av ett av de möjliga svarsalternativen. Svaren ska anges på svarsblanketten (utdelas vid tentamen). Studenter som är godkända på kontrollskrivningen behöver ej besvara uppgift 1–3, utan får tillgodoräkna sig dessa tre uppgifter (i svarsblanketten anges ordet ”Bonus”). Studenter som är godkända på den andra datorlaborationen behöver ej besvara uppgift 12, utan får tillgodoräkna sig denna uppgift (i svarsblanketten anges ordet ”Bonus”). Dessa tillgodoräknanden gäller för den här tentamen och vid omtentamen i april 2024. Gränsen för godkänt är 9 poäng. Möjlighet att komplettera ges för tentander med 8 poäng.

Del II består av uppgifterna 13–16 och varje korrekt lösning ger 10 poäng. Del II rättas bara för studenter som är godkända på eller får komplettera del I och poäng på del II krävs för högre betyg än E. På denna del ska resonemang och uträkningar vara så utförliga och väl motiverade att de är lätta att följa. Införda beteckningar ska förklaras och definieras samt numeriska svar ska anges med minst två värdesiffrors noggrannhet. Studenter som är godkända på den andra datorlaborationen får 3 bonuspoäng på del II (gäller denna tenta och vid omtentamen i april 2024).

Tentamen kommer att vara rättad inom tre arbetsveckor från skrivningstillfället och kommer att finnas tillgänglig på studentexpeditionen minst sju veckor efter skrivningstillfället.

Del I

Uppgift 1

Händelserna F och G är oberoende med $\mathrm{P}(F) = 0.4$ och $\mathrm{P}(G) = 0.7$. Låt $S = F \cup G$.
Bestäm den betingade sannolikheten $\mathrm{P}(F^* \mid S)$.

A: 0.600

B: 0.512

C: 0.420

D: 0.146

Uppgift 2

En sammansatt fyrverkeripjäs (en så kallad ”bomb”) består av tjugo olika raketer som avfyras en efter en. Om en av raketerna är defekt, dvs. om den inte kan avfyras, så avfyras inte de resterande raketerna och hela pjäsen anses vara defekt. Antag att förekomstena av defekter hos raketerna är oberoende och varje raket har sannolikhet 0.01 att vara defekt. Vad är sannolikheten att hela pjäsen är defekt?

A: 0.165

B: 0.200

C: 0.182

D: 0.010

Uppgift 3

De diskreta stokastiska variablerna X och Y har den simultana sannolikhetsfunktionen

$p_{X,Y}(x,y)$	$x=1$	$x=2$
$y=2$	0.4	0
$y=4$	0.5	0.1

Således antar X värdena 1 och 2, medan Y antar värdena 2 och 4.

Beräkna $V(Y-X)$.

A: 0.87

B: 0.89

C: 1.05

D: 1.21

Uppgift 4

Antag att X är $\mathrm{U}(0,16)$-fördelad, att Y är $\mathrm{U}(0,20)$-fördelad och att X och Y är oberoende.

Beräkna $\mathrm{P}\left(\max(X,Y) > 12\right)$.

A: 0.1

B: 0.45

C: 0.55

D: 0.9

Uppgift 5

Låt X och Y vara kontinuerliga och oberoende stokastiska variabler. Deras täthetsfunktioner är $f_X(x)$ respektive $f_Y(y)$, där

$$f_X(x) = \frac{1}{2\sqrt{2\pi}}\, e^{-x^2/8}, \qquad -\infty < x < \infty,$$
$$f_Y(y) = \frac{1}{\sqrt{2\pi}}\, e^{-(y-3)^2/2}, \quad -\infty < y < \infty.$$

Beräkna $\mathrm{P}\,(X \leq Y)$.

A: 0.9099

B: 0.0901

C: 0.5000

D: 0.9582

Uppgift 6

Låt X vara $\mathrm{Bin}(5, 0.9)$-fördelad och Y vara $\mathrm{Bin}(3, 0.9)$-fördelad. Antag att X och Y är oberoende. Beräkna $\mathrm{P}\,(X + Y \leq 6)$.

A: 0.149

B: 0.187

C: 0.750

D: 0.038

Uppgift 7

Låt $x_1 = 15$ och $x_2 = 12$ vara två oberoende observationer från fördelningen $\mathrm{N}(\mu, \sigma)$ och låt $x_3 = 36$ vara en observation, oberoende av de två första, från fördelningen $\mathrm{N}(2\mu, \sigma)$.
Skatta μ med hjälp av minsta-kvadrat-metoden utgående från dessa tre observationer.

A: 11.25

B: 15

C: 15.75

D: 16.5

Uppgift 8

Antag att X är $\mathrm{N}(\mu, \sigma)$-fördelad. Vi tar fram ett stickprov med fem oberoende observationer och får $\bar{x} = 137.0$ och $s^2 = 92.5$, där stickprovsvariansen s^2 är vår skattning av variansen σ^2. Ange det 95%-iga tvåsidiga konfidensintervallet för variansen σ^2.

A: (59.17,125.83)

B: (53.51,131.49)

C: (38.99,521.13)

D: (33.33,770.83)

Uppgift 9

Låt X var en stokastisk variabel med fördelning $\mathrm{Po}(\mu)$. Vi gör 20 oberoende observationer på X och får $\bar{x} = 400$. Ange nedre gränsen för det nedåt begränsade ensidiga konfidensintervallet för μ som har den approximativa konfidensgraden 95%.

A: 360.80

B: 367.10

C: 391.23

D: 392.64

Uppgift 10

Vid en undersökning på ett stort universitet tillfrågades fem slumpvis utvalda studenter om de styrketränade minst två gånger i veckan. Man ville estimera andelen p av alla studenter på universitetet som styrketränade minst två gånger i veckan. För att testa

$$H_0 : p = 0.3$$

mot

$$H_1 : p > 0.3$$

bestämde man sig för beslutsregeln: ”förkasta H_0 om minst fyra av de fem tillfrågade tränar minst två gånger per vecka”. Bestäm testets styrka i punkten $p = 0.8$.

A: ca 0.9997

B: ca 0.82

C: ca 0.74

D: ca 0.03

Uppgift 11

Vid en undersökning rörande trivsel på en mycket stor arbetsplats intervjuades 150 slumpmässigt utvalda män och 100 slumpmässigt valda kvinnor. De fick bland annat svara på frågan: ”Har du någon gång under det senaste året allvarligt funderat på att byta arbete?” Följande svarsfrekvenser erhölls:

	Ja	Nej
Män	63	87
Kvinnor	56	44

Verkar det finnas någon skillnad mellan kvinnor och män i den aktuella frågan?

För att svara på denna fråga utför man ett homogenitetstest, där teststorheten Q fås till $Q = 4.715$. Vilket av följande påståenden är sant?

A: Skillnaden mellan kvinnor och män i den aktuella frågan är varken signifikant på risknivån 2.5% eller risknivån 5%.

B: Skillnaden mellan kvinnor och män i den aktuella frågan är signifikant både på risknivån 2.5% och risknivån 5%.

C: Skillnaden mellan kvinnor och män i den aktuella frågan är signifikant på risknivån 1%, men inte på risknivån 5%.

D: Skillnaden mellan kvinnor och män i den aktuella frågan är signifikant på risknivån 5%, men inte på risknivån 1%.

Uppgift 12

Enligt tidningen *Universitetsläraren* var sjukfrånvaron (mätt i andel av tillgänglig arbetstid) bland KTH:s anställda följande under åren 2012–2022:

År	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Sjukfrånvaro (%)	0.7	1.3	1.9	2.6	2.6	3.0	2.0	1.3	2.0	2.4	2.1

Om vi låter x_i vara antalet år efter 2012 (så att $x_1 = 0$ motsvarar 2012, $x_2 = 1$ motsvarar 2013, osv.) och låter y vara sjukfrånvaron i procent för det året kan vi ta fram summorna

$$\sum_{i=1}^{11} x_i = 55 \quad \sum_{i=1}^{11} x_i^2 = 385 \quad \sum_{i=1}^{11} y_i = 21.9 \quad \sum_{i=1}^{11} y_i^2 = 48.17 \quad \sum_{i=1}^{11} x_i y_i = 118.0.$$

Låt nu nollhypotesen H_0 vara att det inte har skett någon ökning i sjukfrånvaro under de givna åren medan H_1 är att det har skett en sådan ökning. Vilket av följande påståenden stämmer?

A: Nollhypotesen kan inte förkastas varken på signifikansnivå 1% eller 5%.

B: Nollhypotesen kan förkastas på signifikansnivå 5% men inte på signifikansnivå 1%.

C: Nollhypotesen kan förkastas på signifikansnivå 1% men inte på signifikansnivå 5%.

D: Nollhypotesen kan förkastas på både signifikansnivå 1% och 5%.

Del II

Uppgift 13

Frida står i kö för plats på en förskola i Stockholm. Hennes föräldrar kan se att hon är nummer 110 i kön. Antag att barnens antagning till förskolan sker enligt en poissonmodell med intensitetsfaktor 0.4, dvs. att antalet dagar mellan att två intilliggande barn i kön får plats har fördelningen Exp(0.4) och att dessa väntetider är oberoende av varandra.

(a) Beräkna sannolikheten att Frida måste vänta mer än ett år (365 dagar) på att bli antagen till förskolan. Motivera eventuella approximationer och antaganden. (5 p)

(b) Hur många dagar föräldraledighet måste hennes föräldrar ta ut för att försäkra sig om att hon kommer ha en förskoleplats innan slutet av deras ledighet med 95% sannolikhet? (5 p)

Uppgift 14

Den tvådimensionella stokastiska variabeln (X, Y) har täthetsfunktionen

$$f_{X,Y}(x, y) = \begin{cases} e^{-y}, & 0 < x < y < \infty, \\ 0, & \text{för övrigt.} \end{cases}$$

Bestäm $\mathrm{P}(X + Y \geq 1)$. (10 p)

Uppgift 15

(a) Vid mätning av en strömförstärkningsfaktor erhöll man observationerna

23.2 18.3 25.5 26.2 20.4 29.8 29.8 24.8

Hjälpsummor: $\sum_i x_i = 198.0$, $\sum_i x_i^2 = 5017.1$
Mätningarna är behäftade med mätfel varför man antar att data är oberoende observationer från en normalfördelning med väntevärde μ och varians σ^2, båda okända. Ge ett 95%-igt konfidensintervall för μ. (3 p)

(b) Med hjälp av data i del (a), testa på lämpligt sätt hypotesen att $H_0 : \mu = 21.5$ mot $H_1 : \mu \neq 21.5$. (1 p)

(c) En annan person gjorde från 8 andra observationer ett 95%-igt konfidensintervall för μ och erhöll intervallet 23.75 ± 3.87. Slå nu ihop de båda stickproven (dvs. så att alla 16 observationerna får bilda ett stickprov) för att bilda ett nytt 95%-igt konfidensintervall för μ. (5 p)

(d) Med hjälp av konfidensintervallet i del (c), testa på lämpligt sätt hypotesen att $H_0 : \mu = 21.5$ mot $H_1 : \mu \neq 21.5$. (1 p)

Uppgift 16

Låt X vara cauchyfördelad med väntevärde μ, dvs. med täthetsfunktion

$$f_X(x) = \frac{1}{\pi} \cdot \frac{1}{(x-\mu)^2 + 1}.$$

Beräkna ML-skattningen för μ givet observationerna $x_1 = 2.5, x_2 = -0.4$. (10 p)

Lycka till!

Avd. Matematisk statistik

KTH Matematik

LÖSNINGSFÖRSLAG TENTAMEN I SF1917/SF1919 SANNOLIKHETSTEORI OCH STATISTIK, ONSDAG 10 JANUARI 2024 KL 8.00–13.00.

Del I, Svar.

1. B

2. C

3. B

4. C

5. A

6. B

7. D

8. D

9. D

10. C

11. D

12. A

Del I, Lösningsförslag.

Uppgift 1

Rita Venn-diagram. Vi har:

$$P(F^*|S) = P(F^*|(F \cup G)) = \frac{P(F^* \cap (F \cup G))}{P(F \cup G)} = |\text{Venn-diagram}| = \frac{P(F^* \cap G)}{P(F \cup G)} =$$

$$= \frac{P(G) - P(F \cap G)}{P(F) + P(G) - P(F \cap G)} = |\text{oberoendet}| = \frac{P(G) - P(F)P(G)}{P(F) + P(G) - P(F)P(G)} =$$

$$= \frac{0.7 - (0.4)(0.7)}{0.4 + 0.7 - (0.4)(0.7)} = \frac{0.42}{0.82} \simeq 0.512$$

Svaret är B.

Uppgift 2

Låt A_n vara händelsen att den n:te raketen är defekt. Vi har då att $\mathrm{P}(A_n) = 0.01$, vilket ger

$$\mathrm{P}(A_1 \cup A_2 \cup \ldots \cup A_{20}) = 1 - \mathrm{P}(A_1^* \cap A_2^* \cap \ldots \cap A_{20}^*) = 1 - (1 - 0.01)^{20} \approx 0.182.$$

Uppgift 3

$V(Y - X) = C(Y - X, Y - X) = C(Y, Y) + C(Y, -X) + C(-X, Y) + C(-X, -X) =$
$= V(Y) + V(X) - 2C(X, Y)$
$p_X(1) = 0.9, p_X(2) = 0.1$
$E(X) = 1 \cdot 0.9 + 2 \cdot 0.1 = 1.1$
$E(X^2) = 1^2 \cdot 0.9 + 2^2 \cdot 0.1 = 1.3$
$V(X) = E(X^2) - E^2(X) = 1.3 - 1.1^2 = 0.09$
$p_Y(2) = 0.4, p_Y(4) = 0.6$
$E(Y) = 2 \cdot 0.4 + 4 \cdot 0.6 = 3.2$
$E(Y^2) = 2^2 \cdot 0.4 + 4^2 \cdot 0.6 = 11.2$
$V(Y) = E(Y^2) - E^2(Y) = 11.2 - 3.2^2 = 0.96$
$p_{X,Y}(1, 2) = 0.4, p_{X,Y}(1, 4) = 0.5, p_{X,Y}(2, 2) = 0, p_{X,Y}(2, 4) = 0.1$
$E(XY) = 1 \cdot 2 \cdot 0.4 + 1 \cdot 4 \cdot 0.5 + 2 \cdot 2 \cdot 0 + 2 \cdot 4 \cdot 0.1 = 3.6$
$C(X, Y) = E(XY) - E(X)E(Y) = 3.6 - 1.1 \cdot 3.2 = 0.08$
$V(Y - X) = V(Y) + V(X) - 2C(X, Y) = 0.96 + 0.09 - 2 \cdot 0.08 = 0.89$

Uppgift 4

$P(max(X, Y) > 12) = 1 - P(max(X, Y) < 12) = [\text{ober}] =$
$= 1 - P(X < 12) \cdot P(Y < 12) = 1 - \frac{12}{16} \cdot \frac{12}{20} = 1 - \frac{3}{4} \cdot \frac{3}{5} = 0.55$

Uppgift 5

Täthetsfunktioner för X respektive Y ger att:
1) $X \in N(0, 2)$, dvs $\mu_X = 0$, $\sigma_X = 2$,
2) $Y \in N(3, 1)$, dvs $\mu_Y = 3$, $\sigma_Y = 1$.
Det är känd även att X och Y är oberoende.

Följaktligen, $X - Y \in N(0-3, \sqrt{2^2+1^2}) = N(-3, \sqrt{5})$.

Därför, $P(X \leq Y) = P(X - Y \leq 0) =$

$$= \Phi\left(\frac{0-(-3)}{\sqrt{5}}\right) = \Phi\left(\frac{3}{\sqrt{5}}\right) \simeq \Phi(1.34) \simeq |\text{Tabell 1}| \simeq 0.9099$$

Svaret är A.

Uppgift 6

Eftersom $X+Y$ är fördelad enligt $\text{Bin}(5+3, 0.9) = \text{Bin}(8, 0.9)$ har vi

$$\begin{aligned}\mathrm{P}\left(X+Y \leq 6\right) = 1 - \mathrm{P}\left(X+Y > 6\right) = 1 - \mathrm{P}\left(X+Y = 7\right) - \mathrm{P}\left(X+Y=8\right)\\ = 1 - \binom{8}{7} \cdot 0.9^7 \cdot 0.1 - \binom{8}{1} 0.9^8 \approx 0.187.\end{aligned}$$

Uppgift 7

$Q = (x_1-\mu)^2 + (x_2-\mu)^2 + (x_3 - 2\mu)^2$

$$\frac{dQ}{d\mu} = -2(x_1-\mu) - 2(x_2-\mu) - 4(x_3-2\mu) = 0$$

$$\Rightarrow x_1 + x_2 + 2 \cdot x_3 = \mu + \mu + 4 \cdot x_3$$

$$\Rightarrow \mu^*_{obsMK} = \frac{x_1 + x_2 + 2 \cdot x_3}{6} = \frac{15 + 12 + 2 \cdot 36}{6} = 16.5$$

Uppgift 8

Då vi har ett oberoende, normalfördelat stickprov gäller enligt §11.1b) att

$$\frac{(n-1)S^2}{\sigma^2} \in \chi^2(n-1).$$

och enligt §12.4 att

$$f \cdot (\frac{\theta^*}{\theta})^2 \in \chi^2(f).$$

För att skapa det efterfrågade konfidensintervallet, kombinera §12.4 och §11.b.
Då fås konfidensintervallet för σ

$$\left(s\sqrt{\frac{(n-1)}{\chi^2_{0.025}(n-1)}}, s\sqrt{\frac{(n-1)}{\chi^2_{0.975}(n-1)}}\right).$$

Men då vi ska ha konfidensintervallet för σ^2 kvadrerar vi gränserna och får

$$\left(s^2\frac{(n-1)}{\chi^2_{0.025}(n-1)}, s^2\frac{(n-1)}{\chi^2_{0.975}(n-1)}\right).$$

Vilket med insatta numeriska värden ger

$$\left(92.5 \cdot \frac{4}{11.1}, 92.5 \cdot \frac{4}{0.48}\right) = (33.33, 770.83).$$

Uppgift 9

Vi vill använda §12.3 för att kunna bilda ett konfidensintervall med den approximativa konfidensgraden 95% Enligt §12.3 gäller att

$$I_\theta = \theta^*_{obs} \pm D^*_{obs} \cdot \lambda_{\frac{\alpha}{2}}$$

Här är vårt $\theta = \mu$ och $\theta^*_{obs} = \mu^*_{obs} = \bar{x}$. Villkoret enligt §12.3 för att §12.3 ska kunna användas är att $\theta^* \sim$ N-fördelat uppfylls av att $\mu^*_{obs} = \bar{x} = 400 \geq 15$ vilket är villkoret för att approximera en Poissonfördelning till Normalfördelning.

$$V(\bar{X}) = \frac{\sigma^2}{n} = [\text{ty Po-fördelning}] = \frac{\mu}{n} \Rightarrow D^*_{obs} = \sqrt{\frac{\mu^*_{obs}}{n}} = \sqrt{\frac{400}{20}}$$

Nu ska vi ha ett nedåt begränsat konfidensintervall, så vi får

$$I_\theta = (\theta^*_{obs} - D^*_{obs} \cdot \lambda_\alpha, \infty) = (\bar{x} - \sqrt{\frac{\bar{x}}{n}} \cdot \lambda_{0.05}, \infty) = (400 - \sqrt{\frac{400}{20}} \cdot 1.6449, \infty) = (392.64, \infty)$$

Undre gränsen är alltså 392.64.

Uppgift 10

Låt s.v. X ange antal studenter på universitet som styrketränar minst två ggr/v. Då, har vi att $X \in Bin(5, p)$. Styrkan i punkt $p = 0.8$:

$$h(0.8) = \mathrm{P}\left(X \geq 4; p = 0.8\right) = \mathrm{P}\left(X \leq 1; p = 0.2\right) = |\text{Tabell 6 för } n = 5,\ p = 0.2,\ x = 1| = 0.73728$$

Svaret är C.

Uppgift 11

Vi har $(r-1)(c-1) = (2-1)(2-1) = 1$ frihetsgarder. Eftersom teststorheten $Q = 4.715 > 3.84 = \chi^2_{0.05}(1)$ (se Tabell 4, för $\chi^2(1)$-fördelning), så är därför skillnaden mellan män och kvinnor signifikant på risknivån 5%. Dock varken på nivån 2.5% (eftersom $Q = 4.715 < 5.02 = \chi^2_{0.025}(1)$) eller 1% (eftersom $Q = 4.715 < 6.63 = \chi^2_{0.01}(1)$).

Svaret därför är D.

Uppgift 12

Från lydelsen har vi

$$\overline{x} = 5.0 \quad \overline{y} = 1.99091 \Rightarrow s_{xy} = 118.0 - 11 \cdot 5 \cdot 1.99091 \approx 8.49995 \quad s_{xx} = 385 - 11 \cdot 5.0^2 = 110.0$$

vilket ger

$$\beta^\star_{\text{obs}} = \frac{s_{xy}}{s_{xx}} \approx \frac{8.49995}{110.0} \approx 0.07658.$$

Vi har också
$$s_{yy} = 48.17 - 11 \cdot 1.99091^2 \approx 4.56905$$
vilket ger
$$s^2 = \frac{4.56905 - 8.49995^2/110.0}{11-2} = 0.43469.$$
Vi får alltså konfidensintervallen för β till
$$\left[\beta^\star_{\text{obs}} - t_{0.05}(9)\frac{s}{\sqrt{s_{xx}}}, \infty\right) \approx \left[0.07658 - 1.83 \cdot \frac{0.65931}{\sqrt{110.0}}, \infty\right) \approx [-0.03846, \infty)$$
för signifikansnivå 5% och till
$$[-0.10069, \infty)$$
för signifikansnivå 1%.
Eftersom båda konfidensintervall innehåller noll kan vi inte förkasta nollhypotesen varken på signifikansnivå 1% eller 5%. Rätt svar är alltså A.
Notera: Om tvåsidiga konfidensintervall anges får vi samma svar eftersom dessa kommer nödvändigtvis vara ”bredare”.

Del II, Lösningsförslag:

Uppgift 13

Antalet dagar X tills Frida får en plats är alltså summan av 110 stycken oberoende exponentialfördelade stokastiska variabler. Eftersom antalet är stort tillämpar vi centrala gränsvärdessatsen och får att X är approximativt normalfördelat med väntevärde

$$110 \cdot \frac{1}{0.4} = 275$$

samt standardavvikelse

$$\sqrt{110 \cdot \frac{1}{0.4^2}} \approx 26.22$$

(a) Vi får alltså

$$\mathrm{P}\left(X > 365\right) \approx 1 - \Phi\left(\frac{365-275}{26.22}\right) \approx 1 - \Phi(3.43) \approx 1 - 0.9997 = 0.0003$$

(b) Om X är approximativt $\mathrm{N}(275, 26.22)$-fördelat så har vi att

$$\mathrm{P}\left(X > 275 + (26.22) * \lambda_{0.05}\right) \approx \mathrm{P}\left(X > 318.13\right) \approx 0.05.$$

Med andra ord måste de ta ut 319 dagar i föräldraledighet.

Uppgift 14

Vi har:

$$P(X+Y \geq 1) = 1 - P(X+Y<1) = |\text{rita en graf}| = 1 - \int_0^{1/2} dx \left(\int_x^{1-x} e^{-y} dy \right) =$$

$$= 1 - \int_0^{1/2} \left(e^{-x} - e^{x-1} \right) dx = 1 - \Big[-e^{-x} \Big]_0^{1/2} + \Big[e^{x-1} \Big]_0^{1/2} =$$

$$= 1 - \left(-e^{-1/2} + 1 \right) + \left(e^{-1/2} - e^{-1} \right) = 2e^{-1/2} - e^{-1} = 0.845$$

Uppgift 15

a) Ett 95%-igt konfidensintervall ges av $\bar{x} \pm t_{0.025}(n-1)s/\sqrt{n}$, kombinera 11.1 och 12.2 i formelsamlingen. Medelvärdet $\bar{x} = 24.75$ och stickprovsvariansen är $s^2 = \frac{1}{7}(\sum_i x^2 - 8\bar{x}^2) = 116.6/7$. Eftersom $t_{0.025}(7) = 2.36$ erhåller vi intervallet 24.75 ± 3.41.

b) Värdet 21.5 tillhör konfidensintervallet, så hypotesen förkastas ej.

c) Mittpunkten i det andra konfidensintervallet är 23.75. Det medför att medelvärdet för alla 16 observationer är 24.25. Vidare vet vi att $t_{0.025}(7)s_2/\sqrt{8} = 3.87$ (där s_2 är andra stickprovets standardavvikelse) varför $s_2^2 = 3.87^2 \cdot 8/2.36^2 = 21.51$. Eftersom $s_2^2 = \frac{1}{7}(\sum_i y_i^2 - 8\bar{y}^2)$ är $\sum_y y_i^2 = 7s_2^2 + 8\bar{y}^2 = 4663.1$. Det sammanslagna stickprovet får variansen $\frac{1}{15}(\sum_i x_i^2 + \sum_i y_i^2 - 16 \cdot 24.25^2) =$

18.08. Dettta ger ett nytt konfidensintervall $24.25 \pm 2.13\sqrt{18.08/16} = 24.25 \pm 2.26$ eftersom $t_{0.025}(15) = 2.13$.

d) Värdet 21.5 tillhör ej konfidensintervallet, så hypotesen förkastas.

Uppgift 16

Log-likelihoodfunktionen fås till

$$\log L(\mu) = \log\left(\frac{1}{\pi}\frac{1}{(x_1-\mu)^2+1)}\right) + \log\left(\frac{1}{\pi}\frac{1}{(x_2-\mu)^2+1}\right),$$

vilket ger ekvationen

$$(x_1-\mu)((x_2-\mu)^2+1) + (x_2-\mu)((x_1-\mu)^2+1) = 0$$

för de kritiska punkterna. Denna ekvation kan skrivas om till

$$-2\mu^3 + 3(x_1+x_2)\mu^2 - (2 + x_1^2 + x_2^2 + 4x_1x_2)\mu + (x_1 + x_2 + x_1x_2^2 + x_1^2x_2) = 0.$$

Stoppar vi in värdena $x_1 = 2.5$ och $x_2 = -0.4$ får vi

$$-2\mu^3 + 6.3\mu^2 - 4.41\mu = 0,$$

vilket vi kan faktorisera som

$$(-2\mu^2 + 6.3\mu - 4.41)\mu = 0,$$

så vi får rötterna $\mu_1 = 0$, $\mu_2 = 2.1$ och $\mu_3 = 1.05$. Utvärderar vi log-likelihoodfunktionen i dessa punkter har vi

$$\log L(0) \approx -4.4189 \quad \log L(1.05) \approx -4.5539 \quad \log L(2.1) \approx -4.4189.$$

Utöver detta ser vi att $\log L(\mu) \to -\infty$ när $|\mu| \to \infty$, så vi har två kandidater till globala maxima, $\mu_1 = 0$ och $\mu_3 = 2.1$ (μ_2 måste vara ett lokalt minimum eftersom det ligger mellan lokala maxima μ_1 och μ_3).
Vi noterar nu att $\log L(\mu_1) = \log L(\mu_3)$ eftersom $x_1 - \mu_1 = -(x_2 - \mu_3)$ samt $x_2 - \mu_1 = -(x_1 - \mu_3)$ så vi kan skriva

$$\begin{aligned}
\log L(\mu_1) &= \log\left(\frac{1}{\pi}\frac{1}{(x_1-\mu_1)^2+1}\right) + \log\left(\frac{1}{\pi}\frac{1}{(x_2-\mu_1)^2+1}\right)\\
&= \log\left(\frac{1}{\pi}\frac{1}{(x_2-\mu_3)^2+1}\right) + \log\left(\frac{1}{\pi}\frac{1}{(x_1-\mu_3)^2+1}\right)\\
&= \log\left(\frac{1}{\pi}\frac{1}{(x_1-\mu_3)^2+1}\right) + \log\left(\frac{1}{\pi}\frac{1}{(x_2-\mu_3)^2+1}\right) = \log L(\mu_3).
\end{aligned}$$

Vi har då alltså att $\log L(\mu_1) = \log L(\mu_3)$ och därför två globala maxima. ML-skattningen består då av de två värdena 0 och 2.1.