TENTAMEN I SF1920/SF1921 SANNOLIKHETSLÄRA OCH STATISTIK,
MÅNDAG 13 MARS 2023 KL 08.00–13.00

Examinator: Björn-Olof Skytt, tel 08-790 8649.

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. På denna del skall endast svar anges, antingen i form av ett numeriskt värde med tre värdesiffrors noggrannhet eller i form av val av ett av de möjliga svarsalternativen. Svaren anges på svarsblanketten.
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). Detta gäller endast på den här tentan och vid omtentamen i juni 2023. 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 skall resonemang och uträkningar skall vara så utförliga och väl motiverade att de är lätta att följa. Införda beteckningar skall förklaras och definieras och numeriska svar skall anges med minst två värdesiffrors noggrannhet. Studenter som är godkända på den andra datorlaborationen får 3 bonuspoäng på del II på det här ordinarie tentamenstillfället och omtentamenstillfället i juni 2023.

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

För händelserna A och B vet vi att $P(A) = 0.40$, $P(B) = 0.15$ och $P(A^* \cap B) = 0.10$.
Beräkna $P(A \mid B^*)$.

A: $\dfrac{5}{17}$

B: $\dfrac{7}{17}$

C: $\dfrac{10}{17}$

D: $\dfrac{12}{17}$

Uppgift 2

Den stokastiska variabeln $X \in Exp(0.4)$.
$Y = \sqrt{X}$.
Y har täthetsfunktionen $f_Y(y)$. Bestäm $f_Y(2)$.

A: 0.081

B: 0.227

C: 0.323

D: 0.909

Uppgift 3

De stokastiska variablerna Y och Z har den simultana sannolikhetsfunktionen $p_{Y,Z}(y,z) = yz/18, y = 1,2,3, z = 1,2$ Således antar Z alltså värdena 1 eller 2, medan Y antar värdena 1, 2, eller 3. Bestäm $E(YZ)$.

A: 3.00

B: 3.47

C: 3.67

D: 3.89

Uppgift 4

Enligt en undersökning upplever 40% av alla flygledare i landet A oro inför framtiden. Anta att vi helt slumpmässigt väljer ut 12 stycken flygledare i landet A. Vad är sannolikheten att minst 2 av dem känner oro inför framtiden?

A: 0.72

B: 0.77

C: 0.92

D: 0.98

Uppgift 5

Låt X och Y vara stokastiska variabler vars varianser är 4 respektive 9. Om X och Y är beroende, kan variansen av $X+Y$ vara antingen större eller mindre än 13, men den kan inte bli hur stor eller liten som helst. Hur liten kan variansen av $X+Y$ lägst vara, om X och Y är beroende?

A: -59

B: 1

C: 7

D: 11

Uppgift 6

$X \in N(20,\sigma)$.Vi gör 25 oberoende mätningar av X. Hur stort får då σ högst vara för att sannolikheten att medelvärdet $\bar{X}$ ska ligga utanför intervallet (19.95,20.05) ska vara högst 5%?

A: 0.0255

B: 0.0304

C: 0.128

D: 0.152

Uppgift 7

Observationerna x_1, x_2, x_3 är utfall av oberoende och likafördelade stokastiska variabler med väntevärde μ och standardavvikelse 1. För att skatta μ kan man använda sig av
$\hat{\mu}_{obs} = \frac{1}{50}(7x_1 + 21x_2 + 22x_3)$ och $\mu^*_{obs} = \frac{1}{50}(27x_1 + 6x_2 + 17x_3)$.

Vilket av nedanstående påståenden är sant?

A: μ^*_{obs} är den effektivaste skattningen av μ.

B: $\hat{\mu}_{\text{obs}}$ är den effektivaste skattningen av μ.

C: Bägge skattningarna är lika effektiva.

D: Man kan inte avgöra vilken av skattningarna som är effektivast, eftersom minst en av dem inte är väntevärdesriktig.

Uppgift 8

$X \in U(0, \theta)$ och $Y \in U(0, 4\theta)$. X och Y är oberoende. Vi vill skatta θ med Minsta-Kvadrat-metoden och får utfallen x=5.2 och y= 28.3 Bestäm Minsta-Kvadrat-skattningen av θ utgående från dessa data.

A: 12.275

B: 13.4

C: 13.64

D: 13.93

Uppgift 9

Man vill undersöka ifall en viss stråldosmätare A ger högre värden än stråldosmätare B genom att mäta stråldoserna på fem olika platser med väsentligen olika mängder bakgrundstrålning. Observationen på plats i med mätare A, x_i^A, beskrivs av en $\mathrm{N}(\mu_i + \Delta, \sigma)$-fördelad stokastisk variabel. Motsvarande mätning x_i^B med mätare B beskrivs av en $\mathrm{N}(\mu_i, \sigma)$-fördelad stokastisk variabel. Samtliga stokastiska variabler antas vara oberoende.

Plats:	1	2	3	4	5
x_i^A:	5.7	7.0	8.2	4.0	5.8
x_i^B:	6.1	6.3	7.6	4.4	4.9

Tag fram ett tvåsidigt 95% konfidensintervall för Δ och ange övre gränsen.

A: 0.88

B: 1.06

C: 1.95

D: 2.36

Uppgift 10

Antag att $X \in \mathrm{Po}(\mu)$ och låt H_0 vara att $\mu = 1.4$.
Vi vill testa H_0 mot alternativet $H_1 : \mu = 5.0$. Tänk efter om det är för stora eller för små värden på x vi i detta fall förkastar H_0 till förmån för H_1. Vi får en observation : $x = 4$. Bestäm testets p-värde.

A: 0.014

B: 0.054

C: 0.556

D: 0.735

Uppgift 11

Antag att man har stickprov från en fördelning med okänt väntevärde μ. Man önskar testa nollhypotesen H_0: $\mu = 70$ mot H_1: $\mu < 70$. Med hjälp av stickprovet får man fram konfidensintervall för μ med konfidensgraderna 95% och 99%. Konfidensintervallen är antingen

$$A: \quad (-\infty, 69.58) \quad \text{och} \quad (-\infty, 70.44)$$

eller

$$B: \quad (70.42, \infty) \quad \text{och} \quad (69.56, \infty)$$

Vilket av följande påståenden är sant?

A: Med hjälp av intervallen i A ser man att:H_0 kan varken förkastas på risknivån 1% eller risknivån 5%.

B: Med hjälp av intervallen i A ser man att: H_0 kan förkastas både på risknivån 1% och risknivån 5%.

C: Med hjälp av intervallen i A ser man att: H_0 kan förkastas på risknivån 5%, men inte på risknivån 1%

D: Med hjälp av intervallen i B ser man att: H_0 kan varken förkastas på risknivån 1% eller risknivån 5%

E: Med hjälp av intervallen i B ser man att: H_0 kan förkastas både på risknivån 1% och risknivån 5%

F: Med hjälp av intervallen i B ser man att: H_0 kan förkastas på risknivån 5%, men inte på risknivån 1%

Uppgift 12

Följande datamaterial beskriver hur försäljningen av en vara beror av hur mycket som spenderats på reklam.

Reklam (kkr)	22	29	31	35	42
Försäljning (kkr)	105	108	107	100	89

Det är rimligt att tro att det föreligger ett linjärt samband mellan variablerna. Utifrån datamaterialet ovan skattas en linjär regressionsmodell

$$y_i = \alpha + \beta x_i + \varepsilon_i,\ i = 1, \ldots, 5,$$

där y_i = försäljning (kkr) beror av x_i = reklamkostnad (kkr) och ε_i betecknar slumpmässiga fel. Minsta-kvadrat-skattningarna av regressionskoefficienterna α och β blev $\alpha^*_{obs} = 129.3$ respektive $\beta^*_{obs} = -0.865$.
Man kontrollerar vidare om den effekt som reklamkostnaden har på försäljningen är signifikant, dvs man testar $H_0 : \beta = 0$ mot $H_1 : \beta \neq 0$. P-värdet för testet blev 0.089.
Vilken slutsats kan man dra då man fått det P-värdet?

A: Vi kan förkasta H_0 på risknivån 5% och kan därmed förkasta att y_i:na beror av x_i:na på risknivån 5%

B: Vi kan förkasta H_0 på risknivån 5% och kan därmed inte förkasta att y_i:na beror av x_i:na på risknivån 5%

C: Vi kan inte förkasta H_0 på risknivån 5% och kan därmed inte förkasta att y_i:na beror av x_i:na på risknivån 5%

D: Vi kan inte förkasta H_0 på risknivån 5% och kan därmed förkasta att y_i:na beror av x_i:na på risknivån 5%

Var god vänd!

Del II

Uppgift 13

I ett system sitter två komponenter vars livslängder är oberoende stokastiska variabler X och Y, båda med täthetsfunktionen

$$f_X(x) = \begin{cases} \frac{2}{x^3}, & \text{om } \; x \geq 1, \\ 0, & \text{annars,} \end{cases}$$

Systemet brister när båda enheterna har brustit. Låt T vara systemets livslängd. Beräkna väntevärdet $E(T)$ (10 p)

Uppgift 14

En viss typ av mikrokretsar levereras i partier om 1000 enheter. Man använder följande provtagningsschema: Ur varje parti väljs slumpmässigt utan återläggning ett urval om 80 enheter som kontrolleras. Om högst tre enheter i urvalet är defekta godkänns partiet – i annat fall underkänns det. Beräkna

a) sannolikheten att man underkänner ett parti där bråkdelen defekta enheter är 0.005. (5 p)

b) sannolikheten att man godkänner ett parti där bråkdelen defekta enheter är 0.15. (5 p)

Lämpliga och väl motiverade approximationer får göras i både a- och b-delen.

Uppgift 15

Låt X vara $\text{Bin}(n, p)$ och antag att vi har observerat utfallet $x = 1$. Vi vet att n är antingen 2 eller 3, samt att p är antingen 1/2 eller 1/3. Bestäm Maximum-Likelihoodskattningen av n och p.

Ledning: Fundera över innebörden av ML-metoden! (10 p)

Var god vänd!

Uppgift 16

Vid livstidsprovning av elektriska komponenter sätter man n exemplar av komponenten i arbete vid en tidpunkt $t = 0$ och låter dem arbeta under uppsikt tills de upphör att fungera och man registrerar tidpunkterna för detta, dvs livslängderna $x_1, x_2, \cdots, x_n$.

Följande antagande gjordes: Olika exemplars livslängder ses som utfall av oberoende exponentialfördelade stokastiska variabler med väntevärde μ.

Beräkna konfidensintervall för μ med den approximativa konfidensgraden 90% i följande två fall:

a) Man håller kontinuerlig uppsikt och observerar $x_1, x_2, \cdots, x_n$. Ge ett numeriskt svar då $n = 50$ och man observerat livslängderna (ordnade i storleksordning)

0.1	0.1	0.2	0.9	1.2	1.3	1.8	1.8	2.0	2.0
2.0	2.1	2.1	2.3	2.6	2.7	3.0	3.5	3.6	3.8
3.8	3.8	4.8	5.1	5.9	7.3	7.6	7.7	8.5	8.6
8.8	9.1	12.0	12.8	13.4	13.6	14.1	14.2	14.7	16.3
16.7	16.8	16.9	20.6	22.1	25.7	26.3	32.0	33.3	40.3

Som räknehjälp kan meddelas att $\bar{x} = 9.63$. (5 p)

b) Man observerar endast antalet komponenter som fortfarande fungerar vid tiden $t = 6$.
Använd relevanta delar ur ovanstående data för de numeriska beräkningarna. (5 p)

Lycka till!

Avd. Matematisk statistik

KTH Matematik

LÖSNINGSFÖRSLAG
TENTAMEN I SF1920/SF1921 SANNOLIKHETSLÄRA OCH STATISTIK,
MÅNDAG 14 MARS 2022 KL 8.00–13.00.

Del I, Svar

1. B

2. C

3. D

4. D

5. B

6. C

7. B

8. D

9. B

10. B

11. C

12. D

Del I, Lösningsförslag

Uppgift 1

$$P(A \mid B^*) = \frac{P(A \cap B^*)}{P(B^*)} = \frac{P(A) - P(A \cap B)}{1 - P(B)} = \frac{0.40 - 0.05}{1 - 0.15} = \underline{\frac{7}{17}}.$$

Uppgift 2

Ta först fram $F_Y(y) = P(Y \leq y) = P(\sqrt{X} \leq y) = P(X \leq y^2) = F_X(y^2)$
Eftersom $X \in Exp(0.4)$ gäller att

$$f_X(x) = \begin{cases} 0 & \text{om } \; x < 0, \\ 0.4e^{-0.4x} & \text{om } \; x > 0. \end{cases}$$

Då fås $F_X(x) = \int_0^x 0.4e^{-0.4t}dt = 1 - e^{-0.4x}$
Dvs. $F_Y(y) = F_X(y^2) = 1 - e^{-0.4y^2}$
$f_Y(y) = \frac{d}{dy}(1 - e^{-0.4y^2}) = 0.8ye^{-0.4y^2}$
$f_Y(2) = 1.6e^{-1.6} = 0.323$

Uppgift 3

$$E(YZ) = \sum_{alla \;\; y,z} yzp_{Y,Z}(y,z) = 1{\cdot}1\frac{1 \cdot 1}{18} + 2{\cdot}1\frac{2 \cdot 1}{18} + 3{\cdot}1\frac{3 \cdot 1}{18} + 1{\cdot}2\frac{1 \cdot 2}{18} + 2{\cdot}2\frac{2 \cdot 2}{18} + 3{\cdot}2\frac{3 \cdot 2}{18} = \frac{70}{18} = 3.89$$

Uppgift 4

Antal flygledare i vår undersökning som känner oro inför framtiden kan sägas vara en stokastisk variabel X
där $X \in Bin(12, 0.4) = Bin(4, \frac{4}{15})$
$P(X \geq 2) = 1 - P(\leq 1) =$ [se tab 6] $= 1 - 0.0159 = 0.9841 =$ [avrundat] $= 0.98$

Uppgift 5

Eftersom det gäller att korrelationskoefficienten $-1 \leq \rho(X,Y) \leq 1$ och $\rho(X,Y) = \frac{C(X,Y)}{D(X)(D(Y)} = \frac{C(X,Y)}{D(X)(D(Y)} =$ [i vårt fall] $= \frac{C(X,Y)}{2 \cdot 3}$ så måste gälla att $-6 \leq C(X,Y) \leq 6$

$V(X+Y) = V(X) + V(Y) + 2C(X,Y) =$ [i vårt fall] $= 4 + 9 + 2C(X,Y) = 13 + 2C(X,Y)$. Alltså är 1 det minsta värdet $V(X+Y)$ kan anta i vårt fall.

Uppgift 6

$P(19.95 \le \bar{X} \le 20.05) \ge 0.95.$

$$\text{Vi gör om till } N(0,1) \Rightarrow P\left(\frac{19.95 - E(\bar{X})}{D(\bar{X})} \le \frac{\bar{X} - E(\bar{X})}{D(\bar{X})} \le \frac{20.05 - E(\bar{X})}{D(\bar{X})}\right) \ge 0.95$$

$$\Rightarrow P\left(\frac{19.95 - 20}{\sigma/\sqrt{n}} \le \frac{\bar{X} - 20}{\sigma/\sqrt{n}} \le \frac{20.05 - 20}{\sigma/\sqrt{n}}\right) \ge 0.95$$

$$\Rightarrow Y = \frac{\bar{X} - 20}{\sigma/\sqrt{n}}$$

$$\Rightarrow P\left(Y \ge \frac{20.05 - 20}{\sigma/\sqrt{n}}\right) \le 0.025$$

$$\Rightarrow \frac{20.05 - 20}{\sigma/\sqrt{n}} \ge \lambda_{0.025} = 1.96$$

$$\Rightarrow \frac{\sigma}{\sqrt{25}} \le \frac{0.05}{1.96} \Rightarrow \sigma \le 0.128$$

Uppgift 7

1. Eftersom

$$\begin{aligned} E(\hat{\mu}) &= E\left(\frac{1}{50}(7X_1 + 21X_2 + 22X_3)\right) = \frac{1}{50}(7E(X_1) + 21E(X_2) + 22E(X_3) \\ &= \frac{1}{50}(7 + 21 + 22\mu) = \mu \end{aligned}$$

så <u>är $\hat{\mu}_{obs}$ en väntevärdesriktig skattning av μ</u>. Eftersom

$$\begin{aligned} E(\mu^*) &= E\left(\frac{1}{50}(27X_1 + 6X_2 + 17X_3)\right) = \frac{1}{50}(27E(X_1) + 6E(X_2) + 17E(X_3)) \\ &= \frac{1}{50}(27 + 6 + 17\mu) = \mu \end{aligned}$$

så <u>är även μ^*_{obs} en väntevärdesriktig skattning av μ</u>.

2. Nu är

$$\begin{aligned} V(\hat{\mu}) &= V\left(\frac{1}{50}(7X_1 + 21X_2 + 22X_3)\right) = \{\text{ober}\} \\ &= \frac{1}{50^2}(7^2V(X_1) + 21^2V(X_2) + 22^2V(X_3)) = \frac{974}{50^2} \end{aligned}$$

och

$$\begin{aligned} V(\mu^*) &= V\left(\frac{1}{50}(27X_1 + 6X_2 + 17X_3)\right) = \{\text{ober}\} \\ &= \frac{1}{50^2}(27^2V(X_1) + 6^2V(X_2) + 17^2V(X_3)) = \frac{1054}{50^2} > V(\hat{\mu}) \end{aligned}$$

Därför är <u>$\hat{\mu}_{obs}$ en effektivare skattning än μ^*_{obs}</u>.

Uppgift 8

Se §9.2 i F.S. $Q = (x - E(X))^2 + (y - E(Y))^2 =$ Se §4.2 i F.S. $= (x - \frac{\theta}{2})^2 + (y - 2\theta)^2$

$$\frac{dQ}{d\theta} = 0 \Rightarrow 2(x - \frac{\theta}{2})(\frac{-1}{2}) + 2(y - 2\theta)(-2) = 0$$

$$\Rightarrow (x - \frac{\theta}{2}) + 4(y - 2\theta) = 0 \Rightarrow x + 4y = 8\theta + \frac{\theta}{2} = \frac{17}{2}\theta$$

$$\Rightarrow \text{Minsta-Kvadrat.skattningen av } \theta = \frac{2}{17}(x + 4y) = \frac{2}{17}(5.2 + 4 \cdot 28.3) = 13.93$$

Uppgift 9

De parvisa differenserna $z_i = x_i^A - x_i^B$

Plats:	1	2	3	4	5
x_i^A:	5.7	7.0	8.2	4.0	5.8
x_i^B:	6.1	6.3	7.6	4.4	4.9
z:	-0.4	0.7	0.6	-0.4	0.9

sammanfattas av $\overline{z} = 0.28$, $s_z = 0.63$. Nu är $\overline{z}$ ett utfall av $\overline{Z}$ där $\overline{Z}$ är $\mathrm{N}(\Delta, \sigma_z/\sqrt{5})$. Ett 95%-konfidensintervall för Δ ges av

$$\Delta \in \overline{z} \pm t_{0.025}(5-1)\frac{s_z}{\sqrt{5}} = \Delta \in 0.28 \pm 2.78\frac{0.63}{\sqrt{5}} = 0.28 \pm 0.78 = \underline{(-0.50,\ 1.06)}$$

Uppgift 10

Vi förkastar nollhypotesen H_0 till förmån för mothypotesen H_1 om vi får ett stort värde på x i detta fall eftersom $\mu = E(X)$ när $X \in Po(\mu)$. P-värdet är sannolikheten att förkasta nollhypotesen om nollhypotesen är sann utgående från den observation vi fått. I detta fall blr P-värdet $P(X \geq 4)$ om $\mu = 1.4$ Då fås $P(X \geq 4) = 1 - P(X \leq 3) =$ se tab 5 $= 1 - 0.94627 \approx 0.054$.

Uppgift 11

Vi förkastar nollhypotesen H_0 till förmån för mothypotesen H_1 om vi får ett utfall som är tillräckligt mycket mindre än 70 i detta fall. Alltså bildar vi ett ensidigt uppåt begränsat konfidensintervall kring vårt utfall för att se om det täcker över $\mu = 70$ eller inte. Därmed måste intervallen i A användas. Det 99%-ga intervallet måste vara det bredare intervallet av de två och det innesluter $\mu = 70$. Alltså kan vi inte förkasta H_0 på 1 %-nivån, men det 95%-ga intervallet innesluter inte $\mu = 70$, och därmed kan vi förkasta H_0 på 5 %-nivån.

Uppgift 12

P-värdet 0.089 $>$ risknivån $\alpha = 0.05$,så vi kan inte förkasta H_0 att $\beta = 0$ på 5%-nivån. Dämed kan vi förkasta att y_i:na beror av x_i:na på risknivån 5%.

Del II, Lösningsförslag

Uppgift 13

Systemets livslängd är $T = \max(X, Y)$. Vi får därför att

$$F_T(x) = P(T \leq x) = P(X \leq x, Y \leq x) = P(Y \leq x)P(Y \leq x) = F(x)^2$$

där $F(x)$ är fördelningsfunktionen till X och Y. Vi har att

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{1}^{x} \frac{2}{t^3}dt = 1 - \frac{1}{x^2} \quad \text{för } x \geq 1$$

Fördelningsfunktionen för T blir alltså $F_T(x) = F(x)^2$ och täthetsfunktionen fås genom derivering;

$$f_T(x) = 2F(x)f(x) = 2(1 - \frac{1}{x^2})\frac{2}{x^3} = \frac{4}{x^3} - \frac{4}{x^5} \quad \text{för } x \geq 1$$

.
Härav erhålles

$$E(T) = \int_{-\infty}^{\infty} x f_T(x)dx = \int_{1}^{\infty} (\frac{4}{x^2} - \frac{4}{x^4})dx = \left[-\frac{4}{x} + \frac{4}{3x^3} \right]_1^{\infty} = \underline{2\frac{2}{3} \approx 2.67}.$$

Uppgift 14

X = antalet defekta i urvalet.
a) X är Hyp$(1000, 80, 0.005)$-fördelad. Då $n/N = 80/1000 = 0.08 \leq 0.1$ så är X approximativt Bin$(80, 0.005)$-fördelad.
Då $p = 0.005 \leq 0.1$ så är X approximativt Po$(80 \cdot 0.005)$ =Po(0.4)-fördelad. Således fås
$P(X > 3) = 1 - P(X \leq 3) =$ (Poisson-tabellen) $\approx 1 - 0.99922 = \underline{0.00078}$.

b) X är Hyp$(1000, 80, 0.15)$-fördelad. Som ovan fås att X approximativt Bin$(80, 0.15)$-fördelad.
Då $np(1-p) = 80 \cdot 0.15 \cdot 0.85 = 10.2 \geq 10$ så är X approximativt $N(80 \cdot 0.15, \sqrt{10.2}) = N(12, \sqrt{10.2})$-fördelad. Således fås

$$P(X > 3) = P\left(\frac{X - 12}{\sqrt{10.2}} > \frac{3 - 12}{\sqrt{10.2}}\right) = 1 - \Phi\left(-\frac{9}{\sqrt{10.2}}\right) = \Phi\left(\frac{9}{\sqrt{10.2}}\right) = \Phi(2.82) = \underline{0.9976}.$$

Uppgift 15

Inför parametern $\theta = (n,p)$ som alltså har 4 tänkbara värden $(2,1/3),(2,1/2),(3,1/3)$ samt $(3,1/3)$. ML-metoden innebär att vi som skattning av parametern θ tar det θ-värde θ^* som gör sannolikheten för det observerade utfallet (dvs här $x=1$) så stor som möjligt. Vi får lätt

$$P(X=1;\theta=(2,1/3)) = \binom{2}{1}\left(\frac{1}{3}\right)^1\left(1-\frac{1}{3}\right)^{2-1} = \frac{4}{9}$$

$$P(X=1;\theta=(2,1/2)) = \binom{2}{1}\left(\frac{1}{2}\right)^1\left(1-\frac{1}{2}\right)^{2-1} = \frac{1}{2}$$

$$P(X=1;\theta=(3,1/3)) = \binom{3}{1}\left(\frac{1}{3}\right)^1\left(1-\frac{1}{3}\right)^{3-1} = \frac{4}{9}$$

$$P(X=1;\theta=(3,1/2)) = \binom{3}{1}\left(\frac{1}{2}\right)^1\left(1-\frac{1}{2}\right)^{3-1} = \frac{3}{8}$$

Eftersom 1/2 är det största av dessa 4 värden är alltså $\underline{n=2, p=1/2}$ ML-skattningen av $\theta = (n,p)$.

Uppgift 16

a) Vi skattar μ med $\mu^* = \bar{x} = \frac{1}{50}\sum_{i=1}^{50} x_i = 9.63$. Motsvarande stickprovsvariabel $\bar{X}$ är approximativt $N(\mu, \mu/\sqrt{50})$ enligt CGS eftersom $E(X)=\mu$ och $V(X)=\mu^2$. Vi får då det approximativt 90%-iga konfidensintervallet

$$\bar{x} \pm \lambda_{0.05}\frac{\bar{x}}{\sqrt{50}} \approx 9.63 \pm 1.6449\frac{9.63}{\sqrt{50}} \approx \underline{9.63 \pm 2.24 = (7.39, 11.87)}.$$

b) Låt Y = antalet komponenter som fungerar vid tiden $t=6$. Y är $\text{Bin}(50,p)$ där

$p = P(X \geq 6) = e^{-6/\mu}$. Vi får eftersom 25 st av observationerna är mindre än 6 att $p^*_{obs} = 25/50 = 1/2$ och $np(1-p) \approx np^*_{obs}(1-p^*_{obs}) = 12.5 > 10$ och således är normalapproximation av $\text{Bin}(50,p)$ tillåten, dvs $\text{Bin}(50,p) \approx N(50p, \sqrt{50p(1-p)})$. Detta ger att p^*_{obs} är approximativt $N(p, \sqrt{p(1-p)/50})$.
Alltså är ett approximativt 90%-igt konfidensintervall för p

$$p^*_{obs} \pm \lambda_{0.05}\sqrt{\frac{p^*_{obs}(1-p^*_{obs})}{50}} \approx 0.5 \pm 0.117 = (0.383, 0.617).$$

Eftersom $p = e^{-6/\mu}$ dvs $\mu = -6/\ln p$ blir konfidensintervallet för μ

$$\left(-\frac{6}{\ln 0.383}, -\frac{6}{\ln 0.617}\right) \approx \underline{(6.25, 12.43)}.$$

Man kan notera att konfidensintervallet i b-delen blir lite bredare och detta beror ju på att vi i b-delen utnyttjar mindre av informationen i stickprovet än i a-delen.