Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF1920/SF1921 SANNOLIKHETSTEORI OCH STATISTIK,
TISDAG 10 MARS 2020 KL 8.00–13.00.

Examinator för SF1920/SF1921: Björn-Olof Skytt, 08-790 86 49.

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. Studenter som är godkända på kontrollskrivningen behöver ej besvara uppgift 1-3 på del I, utan får tillgodoräkna sig dessa tre uppgifter av den ordinarie tentamen och den första omtentamen. Gränsen för godkänt är 9 poäng. Möjlighet att komplettera ges för tentander med 8 poäng. Tid och plats för komplettering kommer att anges på kursens hemsida.
Del II består av uppgifterna 13-16 och varje korrekt lösning ger 10 poäng. Del II rättas bara för studenter med minst 8 poäng på del I och poäng på del II krävs för högre betyg än E. På denna del skall resonemang och uträkningar 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 datorlaboration 2 får tillgodoräkna sig uppgift 12 på del I och får dessutom 3 bonuspoäng på del II av den ordinarie tentamen och den första omtentamen.

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 gäller att ett av de fem påståendena nedan är felaktigt. Vilket?

A: $P(A \cup B) = P(A) + P(B \cap A^*)$

B: $P(A \cap B) = P(B \mid A) \cdot P(A)$

C: $P(A \cap B^*) = P(A) - P(B \cap A)$

D: $P(B \mid A) = 1 - P(B^* \mid A^*)$

E: $P(B) = P(B \cap A) + P(B \cap A^*)$

Uppgift 2

En stokastisk variabel X har fördelningsfunktionen

$$F_X(x) = \begin{cases} 1 - e^{-x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Bestäm $E\big(e^{-X}\big)$.

Uppgift 3

Den tvådimensionella s.v. (X, Y) antas vara likformigt fördelad på kvadraten
$K = \{(x, y) : -1 \leq x \leq 1, -1 \leq y \leq 1\}$. Beräkna $P(0 \leq X + Y \leq 1.5)$.

A: 0.125

B: 0.438

C: 0.469

D: 0.500

Uppgift 4

En basketspelare skjuter 12 straffkast. Antag att sannolikheten att hen sätter ett straffkast är 0.8 och att sannolikheten inte förändras för olika kast. Antag också att händelserna att sätta ett straffkast är oberoende av varandra. Bestäm sannolikheten att antalet satta straffkast överstiger 7 men ej 10.

A: 0.520

B: 0.653

C: 0.706

D: 0.859

Uppgift 5

Sara ska veckohandla i en stor affär och göra ett inköp på 48 artiklar. Antag att alla belopp avrundas vart och ett till hela kronor. Avrundningsfelen $X_1, X_2, \ldots, X_{48}$ antas vara oberoende stokastiska variabler som är likformig fördelade $U(-0.5, 0.5)$. Saras totala avrundningsfel betecknas $X = X_1 + X_2 + \ldots + X_{48}$. Beräkna den approximativa sannolikheten $P(-2 \leq X \leq 2)$.

A: 0.046

B: 0.317

C: 0.683

D: 0.954

Uppgift 6

Låt X och Y vara två oberoende stokastiska variabler sådana att $E(X) = E(Y) = 1$, $D(X) = 11$ och $C(XY, Y) = 9$. Bestäm $D(Y)$.

Uppgift 7

Låt X_1, X_2 och X_3 vara tre oberoende likafördelade stokastiska variabler sådana att $X_i \in \mathrm{Po}\,(\mu)$. Beräkna maximum-likelihood-skattningen av μ då $x_1 = 4$, $x_2 = 10$ och $x_3 = 1$.

Uppgift 8

Antag att vi för att skatta väntevärdet μ i en normalfördelning, med okänd standardavvikelse σ, planerar att ta ett slumpmässigt stickprov $x_1, x_2, \ldots, x_n$ av storlek $n = 15$. Medelvärdet blev $\overline{x} = 20$ och standardavvikelsen $s = 5$. Då blir ett 95% konfidensintervall för väntevärdet μ:

A: $(17.23, 22.77)$

B: $(17.42, 22.58)$

C: $(17.73, 22.27)$

D: $(18.71, 21.29)$

Uppgift 9

Antag att vi gör två stickprov $x_1, x_2, \ldots, x_n$ och $y_1, y_2, \ldots, y_m$ från två oberoende normalfördelade populationer $N(\mu_1, \sigma_1)$ respektive $N(\mu_2, \sigma_2)$. Stickprovet $x_1, x_2, \ldots, x_n$ gav medelvärdet $\overline{x}$ och standardavvikelsen s_x, medan stickprovet $y_1, y_2, \ldots, y_m$ gav medelvärdet $\overline{y}$ och standardavvikelsen s_y. Antag vidare att i populationen $\sigma_1 = \sigma_2$ och att de är okända. Vi testar $H_0 : \mu_1 = \mu_2$ mot $H_1 : \mu_1 < \mu_2$, på signifikans nivån 1%. Som testvariabel används: $t = \frac{\overline{x}-\overline{y}}{s\sqrt{\frac{1}{n}+\frac{1}{m}}}$, där $s^2 = \frac{(n-1)s_x^2+(m-1)s_y^2}{n+m-2}$. Om stickprovsstorlekarna är $n = 12$ och $m = 13$, så förkastas H_0 om

A: $t < -2.81$

B: $|t| > 2.81$

C: $t < -2.50$

D: $|t| > 2.50$

Uppgift 10

Antag att ett parti fick 17% av rösterna i senaste val. Enligt en gallupundesökning omfattande 900 personer har andelen vid ett senare tillfälle skattats till 18.5%. Kan vi nu dra slutsatsen att stödet har ökat? Svara på frågan genom att ange P-värdet för det motsvarande testet.
Ledning: Använd approximationen $Bin(n, p) \sim N(np, \sqrt{np(1-p)})$ om $np(1-p) \geq 10$.

A: 0.058

B: 0.115

C: 0.246

D: 0.877

Uppgift 11

Biverkningseffekter av två typer av medicin, benämnda Typ 1 och Typ 2, har studerats i en grupp om 570 personer. Resultatet sammanfattas i tabellen nedan.

	Medicin av Typ 1	Placebo	Medicin av Typ 2
Antal som upplevde biverkningar	27	10	17
Antal som inte upplevde biverkningar	373	70	73

För att testa nollhypotesen H_0 : Det finns ingen skillnad mellan grupperna, beräknar man teststorheten Q och får $Q = 13.62$. Vilken slutsats kan man dra då man fått denna teststorhet?

A: H_0 kan varken förkastas på risknivån 1% eller risknivån 0.1%

B: H_0 kan förkastas på risknivån 1%, men inte på risknivån 0.1%

C: H_0 kan förkastas på risknivån 0.1%, men inte på risknivån 1%

D: P-värdet är större än 0.05

Uppgift 12

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

Reklamkostnad (kkr)	12	19	21	25	28
Försäljning (kkr)	96	108	114	110	127

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

$$y_i = \alpha + \beta x_i + \varepsilon_i,$$

där y_i = försäljning (kkr) beror av x_i = reklamkostnad (kkr) och ε_i betecknar slumpmässiga fel, $i = 1, \ldots, 5$. Bestäm minsta-kvadrat-skattningarna α^*_{obs} och β^*_{obs} av regressionskoefficienterna α respektive β.

A: $\alpha^*_{obs} = 1.66$, $\beta^*_{obs} = 76.14$

B: $\alpha^*_{obs} = 1.66$, $\beta^*_{obs} = 145.86$

C: $\alpha^*_{obs} = 76.14$, $\beta^*_{obs} = 1.66$

D: $\alpha^*_{obs} = 145.86$, $\beta^*_{obs} = 1.66$

Del II

Uppgift 13

Felkvoten vid en tillverkningsprocess är 0.03. Ur ett parti om 5000 tillverkade enheter väljer man på måfå 100 enheter och justerar processen om fler än k av dessa är defekta. Låt Y beteckna antalet defekta enheter bland dom 100 utvalda. Bestäm det minsta heltal k, för vilket det gäller att $P(\text{processen justeras}) = P(Y > k) < 0.01$. *Ledning:* Välmotiverade approximationer är tillåtna. (10 p)

Uppgift 14

Anna köper en utomhusbelysning. Belysningen består av två oberoende (i statistisk mening) parallellkopplade lampor. Hon vet av erfarenhet att livslängden för var och en av lamporna är exponentialfördelad med parametern λ. Livslängden T för belysningen är då livslängden för den lampa som går sönder *sist*.

a) Vilken fördelning har livslängden T av utomhusbelysningen? Svara på frågan genom att ange täthetsfunktionen $f_T(t)$ för T. *Ledning:* ange först fördelningsfunktionen $F_T(t)$ för T. (6 p)

b) Bestäm den förväntade livslängden $E(T)$ av belysningen som en funktion av λ. (4 p)

Uppgift 15

Om två oberoende exponentialföredelade stokastiska variabler med parameter a adderas får man en s.k. Erlangfördelning, efter en dansk teleingenjör. Denna fördelning har viktiga tekniska tillämpningar: t.ex. är summan av väntetid och betjäningstid i ett kösystem ofta Erlangfördelad, eftersom man antar att såväl väntetid som betjäningstid är exponentialfördelade. Erlangfördelningen har täthetsfunktionen

$$f_X(x) = \begin{cases} a^2 x e^{-ax}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

a) Bestäm ML-skattningen av parametern a baserad på obersvationerna $x_1, x_2, \ldots, x_n$ av den Erlangfördelade stokastiska variabeln X. (5 p)

b) Bestäm MK-skattningen av parametern a baserad på obersvationerna $x_1, x_2, \ldots, x_n$ av den Erlangfördelade stokastiska variabeln X. *Ledning:* bestäm först väntevärdet $E(X)$. (5 p)

Uppgift 16

Försökpersoners reaktionstider i sekunder vid ett visst psykologiskt test kan antas vara normalfördelade med väntevärde μ och den kända standardavvikelsen $\sigma = 2$. För att testa $H_0 : \mu = 25$ mot $H_1 : \mu < 25$ gör man n st observationer och får medelvärdet $\bar{x}$.

a) Ställ upp det villkor, uttryckt i $\bar{x}$ och n, som skall gälla för att H_0 skall förkastas på 1%-nivån. (4 p)

b) Hur många obersvationer n behövs om man vill att styrkan för alternativet $\mu = 23$ skall vara minst 0.98 ? (6 p)

Lycka till!

LÖSNINGSFÖRSLAG TENTAMEN I SF1920/SF1921 SANNOLIKHETSTEORI OCH STATISTIK, TISDAG 10 MARS 2020 KL 8.00–13.00.

Del I, Svar

1. D
2. 0.5
3. C
4. B
5. C
6. 3
7. 5
8. A
9. C
10. B
11. B
12. C

Del II, Svar

13. $k = 8$

14a. Täthetsfunktionen $f_T(t) = 2\lambda(1 - e^{-\lambda t})e^{-\lambda t}$
14b. $E(T) = \frac{3}{2\lambda}$

15a. $a^*_{ML,obs} = 2/\bar{x}$
15b. $a^*_{MK,obs} = 2/\bar{x}$

16a. Förkasta H_0 om $\bar{x} < 25 - \lambda_{0.01}\frac{2}{\sqrt{n}}$, där $\lambda_{0.01} = 2.3263$ (Tabell 2)
16b. Minst $n = 20$ observationer

LÖSNINGSFÖRSLAG TENTAMEN I SF1920/SF1921 SANNOLIKHETSTEORI OCH STATISTIK, TISDAG 10 MARS 2020 KL 8.00–13.00.

Del I, Lösningar

Uppgift 1

Svar: D. Det korrekta uttrycket är: $P(B|A) = 1 - P(B^*|A)$

Uppgift 2

Vi deriverar $F_X(x)$ för att få täthetsfunktionen $f_X(x)$. Därmed

$$E\big(e^{-X}\big) = \int_0^\infty e^{-x} f_X(x)\mathrm{d}x = \int_0^\infty e^{-x}e^{-x}\mathrm{d}x = \int_0^\infty e^{-2x}\mathrm{d}x = \big(-\frac{1}{2}e^{-2x}\big)\Big|_0^\infty = 0 - \big(-\frac{1}{2}e^0\big) = \frac{1}{2}$$

Svar: 0.5

Uppgift 3

Genom att tänka geometriskt/grafiskt, för en tvådimensionell s.v. (X, Y) som är likformigt fördelad på kvadraten $K = \{(x, y) : -1 \leq x \leq 1, -1 \leq y \leq 1\}$:

$$P(0 \leq X + Y \leq 1.5) = \frac{\text{arean av } \{(x, y) : -x \leq y \leq -x + 1.5\}}{\text{arean av } \{K\}} = \big|\text{rita}\big| = \frac{\frac{2\cdot 2}{2} - \frac{(0.5)(0.5)}{2}}{2\cdot 2} = 0.4688$$

Svar: C

Uppgift 4

Låt s.v. Y vara antal straffkast av 12 hen *inte* sätter. Då har vi att $Y \in Bin(12, 0.2)$, och det vi söker därmed är

$$P(2 \leq Y \leq 4) = P(Y \leq 4) - P(Y \leq 1) = |\text{Tabell } 6| = 0.92744 - 0.27488 = 0.65256$$

Svar: B

Uppgift 5

Väntevärdet μ och standardavvikelsen σ för $X_i \in U(-0.5, 0.5)$ är $\mu = \frac{-0.5+0.5}{2} = 0$ resp $\sigma = \sqrt{\frac{0.5-(-0.5)}{12}} = \sqrt{\frac{1}{12}}$, $i = 1, \ldots, 48$. Vidare, enligt centrala gränsvärdesatsen (CGS) är $X = X_1 + \ldots + X_{48}$ approximativt $N(0\cdot 48, \sqrt{\frac{1}{12}}\cdot\sqrt{48}) = N(0, 2)$. Således,

$$P(-2 \leq X \leq 2) = P(X \leq 2) - P(X \leq -2) = \Phi\Big(\frac{2-0}{2}\Big) - \Phi\Big(\frac{-2-0}{2}\Big) = \Phi(1) - \Phi(-1) =$$

$$= \Phi(1) - (1 - \Phi(1)) = 2\Phi(1) - 1 = |\text{Tabell } 1| = 2(0.8413) - 1 = 0.6826.$$

Svar: C

Uppgift 6

Vi har $C(XY,Y) = E(X \cdot Y^2) - E(XY) \cdot E(Y) = |\text{X och Y är oberoende}| = E(X) \cdot E(Y^2) - E(X) \cdot \big(E(Y)\big)^2$. Vidare, eftersom $E(X) = 1$ och $C(XY,Y) = 9$,

$$D(Y) = \sqrt{V(Y)} = \sqrt{E(Y^2) - (E(Y))^2} = \sqrt{C(XY,Y)} = \sqrt{9} = 3$$

Svar: 3

Uppgift 7

Först ställer vi upp likelihood-funktionen. Eftersom $X_i \in Po(\mu), i = 1,2,3$, har vi

$$L(\mu) = \prod_{i=1}^{3} \frac{\mu^{x_i}}{(x_i)!} e^{-\mu} = \frac{\mu^{\sum_i x_i}}{(x_1)!(x_2)!(x_3)!} \cdot e^{-3\mu}.$$

Vi deriverar log-likelihooden $L(\mu)$ och löser ekvationen $\frac{\mathrm{d}\ln(L(\mu))}{\mathrm{d}\mu} = 0$:

$$\frac{\mathrm{d}\ln(L(\mu))}{\mathrm{d}\mu} = \frac{\mathrm{d}}{\mathrm{d}\mu}\Big(\sum_i x_i \ln(\mu) - \ln((x_1)!(x_2)!(x_3)!) - 3\mu\Big) = 0,$$

$$\frac{\sum_i x_i}{\mu} - 3 = 0.$$

Således, maximum-likelihood-skattningen av μ är $\mu^*_{obs} = \frac{\sum_i x_i}{3} = \frac{4+10+1}{3} = 5$.
Svar: 5

Uppgift 8

Vi har, eftesom σ är okänd,

$$\begin{aligned} I_\mu(0.95) &= \Big(\bar{x} \pm t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\Big) = \Big|t_{\alpha/2}(n-1) = t_{0.025}(14) = 2.14, \text{Tabell 3}\Big| \\ &= \Big(20 \pm 2.14\frac{5}{\sqrt{15}}\Big) = \big(20.0 \pm 2.7627\big) \\ &= (17.2373, 22.7627) = |\text{Gardering utåt för minst 95\% konfidens}| \\ &= (17.23, 22.77). \end{aligned}$$

Svar: A

Uppgift 9

Vi vet att $\sigma_1 = \sigma_2$ är okända, $n = 12$ och $m = 13$, samt $\alpha = 0.01$.
Testvariabeln $t = \frac{\bar{x}-\bar{y}}{s\sqrt{\frac{1}{n}+\frac{1}{m}}}$ är $t(n+m-2)$ fördelad, där $s^2 = \frac{(n-1)s_x^2+(m-1)s_y^2}{n+m-2}$.
Således, eftersom $H_1 : \mu_1 < \mu_2$, förkastas nollhypotesen om

$$t < -t_\alpha(n+m-2),$$

där $t_\alpha(n+m-2) = t_{0.01}(23) = 2.5$, från Tabell 3.
Svar: C

Uppgift 10

Vi testar $H_0 : p = 0.17$ mot $H_1 : p > 0.17$ i $X \in Bin(n, p)$-fördelning, där X är antal av $n = 900$ undersökta som stödjer partiet. ML-skattingen av p är $p^* = X/n$. Denna skattning/andelen är även väntevärdesriktig ($E(p^*) = p$) och därför används som en teststorhet för att pröva $H_0 : p = 0.17$ mot $H_1 : p > 0.17$. Om H_0 är sann, gäller att $np(1-p) = 900(0.17)(1-0.17) = 127 > 10$, och därför enligt approximationen

$$X \in Bin(n, p) \simeq N(np, \sqrt{np(1-p)}).$$

Storheten $p^* = X/n$ är således också approx normalfördelad:

$$p^* = \frac{X}{n} \simeq N\left(\frac{np}{n}, \sqrt{\frac{np(1-p)}{n^2}}\right) = N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

om H_0 är sann. Enligt galuppundersökning, fick vi $p^*_{obs} = 0.185$, och därför har vi för $H_0 : p = 0.17$ mot $H_1 : p > 0.17$:

$$\begin{aligned} \text{P-värdet} &= P\big(p^* \geq p^*_{obs} | H_0 \text{ är sann}\big) = P\left(p^* \geq 0.185 \middle| p^* \simeq N\Big(0.17, \sqrt{0.17(1-0.17)/900}\Big)\right) \\ &= 1 - \Phi\left(\frac{0.185 - 0.17}{\sqrt{\frac{0.17(1-0.17)}{900}}}\right) = 1 - \Phi(1.2) = |\text{Tabell 1}| = 1 - 0.885 = 0.115 \end{aligned}$$

Svar: B

Uppgift 11

H_0 förkastas på risknivån α, om teststorheten $Q > \chi^2_\alpha(2) = \chi^2_\alpha((2-1)(3-1))$.
Från Tabell 4, $\chi^2_{0.01}(2) = 9.21$ och $\chi^2_{0.001}(2) = 13.8$. Eftersom $Q = 13.62$, innebär detta att H_0 kan förkastas på risknivån 1%, men inte på risknivån 0.1%.
OBS! P-värdet för testet kommer att vara mindre än 0.01, eftersom $Q = 13.62 > 9.21 = \chi^2_{0.01}(2)$, och således kan inte vara större än 0.05.
Svar: B

Uppgift 12

Alternativen A och B kullkastas på en gång. Detta eftersom, om man räknar på varje av datapunkter, får vi inte ihop något vettigt alls. Exempelvis, för elternativet A och den första datapunkten, har vi: $\alpha^*_{obs} + \beta^*_{obs} x_1 = 1.66 + (76.14)(12) = 915.34$ jmf med 96 som är y_1.
Med samma resonemang kan man lätt utesluta även det svarsalternativet D. Alltså, den korrekta MK-modellen blir: $\hat{y}_i = 76.14 + 1.66\, x_i$, alternativ C.
Man kan även använda sig av miniräknare eller formelsamling för att beräkna α^*_{obs} och β^*_{obs}. Svaret blir detsamma.
Svar: C

Del II, Lösningar

Uppgift 13

Stokastisk variabel $Y \in Hyp(N, n, p)$, där $N = 5000, n = 100, p = 0.03$. Eftersom $n/N = 100/5000 = 0.02 < 0.1$, följande approximation kan användas:

$$Y \in Hyp(N, n, p) \sim Bin(n, p).$$

Och sedan, eftersom $p = 0.03 < 0.1$, så approximerar vi Bin-förd med Po-förd:

$$Y \sim Bin(n, p) \sim Po(np) = Po(3).$$

Från Tabell 5 för $Y \sim Po(3)$ ser vi att det minsta heltal $k = 8$, för vilket

$$P(\text{processen justeras}) = P(Y > 8) = 1 - P(Y \leq 8) = 1 - 0.9962 < 0.01.$$

Detta eftersom $P(Y > 7) = 1 - P(Y \leq 7) = 1 - 0.9881 > 0.01$, och även för alla $m < 7$ på samma sätt kan man visa att $P(Y > m) > 0.01$; men $P(Y > 9) = 1 - P(Y \leq 9) = 1 - 0.9989 < 0.01$, och även för alla $l > 9$ på samma sätt kan man visa att $P(Y > l) < 0.01$.
Svar: $k = 8$

Uppgift 14

Låt ξ_i vara livslängd för lampa i, $\xi_i \in \text{Exp}(\lambda)$, $i = 1, 2$. Dvs. $\xi_i, i = 1, 2$, är oberoende likafördelade slumpvariabler med densamma fördelningsfunktion $F_{\xi_1}(t) = 1 - e^{-\lambda t}, t \geq 0$.
Slumpvariabel T, livslängden av den lampan som går sönder *sist*, definieras i så fall på följande sätt:

$$T = \max(\xi_1, \xi_2).$$

a) Fördelsningsfunktionen $F_T(t)$ för T blir då:

$$\begin{aligned} F_T(t) = P(T \leq t) &= P(\max(\xi_1, \xi_2) \leq t) \\ &= P(\{\xi_1 \leq t\} \cap \{\xi_2 \leq t\}) = |\xi_1 \text{ och } \xi_2 \text{ är oberoende}| \\ &= \prod_{i=1}^{2} P(\xi_i \leq t) = \big(F_{\xi_1}(t)\big)^2 \\ &= \big(1 - e^{-\lambda t}\big)^2,\ t \geq 0. \end{aligned}$$

Täthetsfunktionen $f_T(t)$ för T får vi sedan genom att derivera $F_T(t)$:

$$f_T(t) = \frac{\mathrm{d}F_T(t)}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\big(1 - e^{-\lambda t}\big)^2\Big) = 2\big(1 - e^{-\lambda t}\big)\lambda e^{-\lambda t}.$$

Svar: $f_T(t) = 2\lambda\big(1 - e^{-\lambda t}\big)e^{-\lambda t}$

b) Vi har:

$$\begin{aligned}
E(T) &= \int_0^{\infty} t f_T(t)\mathrm{d}t \\
&= \int_0^{\infty} t2\lambda\big(1-e^{-\lambda t}\big)e^{-\lambda t}\mathrm{d}t \\
&= 2\int_0^{\infty} t\Big(\lambda e^{-\lambda t}\Big)\mathrm{d}t \;-\; \int_0^{\infty} t\Big(2\lambda e^{-2\lambda t}\Big)\mathrm{d}t \\
&= \Big|E(Y) = \int_0^{\infty} y\Big(\mu e^{-\mu y}\Big)\mathrm{d}y = \frac{1}{\mu},\ \text{om } f_Y(y) = \mu e^{-\mu y}, y \geq 0,\ \text{dvs } Y \in \mathrm{Exp}(\mu)\Big| \\
&= 2\frac{1}{\lambda} - \frac{1}{2\lambda} = \frac{3}{2\lambda}
\end{aligned}$$

Svar: $E(T) = \frac{3}{2\lambda}$

Uppgift 15

a) Antag att $a > 0$. Först sätter vi upp likelihoodfunktionen:

$$\begin{aligned}
L(a) = f(x_1, \ldots, x_n; a) = |\text{oberoendet}| &= \prod_{i=1}^{n} f(x_i; a) \\
= \prod_{i=1}^{n} a^2 x_i e^{-a x_i} &= a^{2n} \cdot \big(\prod_{i=1}^{n} x_i\big) \cdot e^{-a\sum_{i=1}^{n} x_i}
\end{aligned}$$

Efter det logaritmeras likelihoodfunktionen $L(a)$:

$$\begin{aligned}
\ln L(a) &= \ln\Big(a^{2n} \cdot \big(\prod_{i=1}^{n} x_i\big) \cdot e^{-a\sum_{i=1}^{n} x_i}\Big) \\
&= 2n\ln(a) + \sum_{i=1}^{n} \ln(x_i) - a\sum_{i=1}^{n} x_i
\end{aligned}$$

För att leta reda på a som maximerar $\ln L(a)$, deriverar man först funktionen $\ln L(a)$ m.a.p. på a. Sedan sätter man upp följande ekvation och löser den för a:

$$\begin{aligned}
\frac{\mathrm{d}\ln L(a)}{\mathrm{d}a} &= 0,\ \text{som blir} \\
\frac{2n}{a} - \sum_{i=1}^{n} x_i &= 0.
\end{aligned}$$

Således, blir ML-skattningen av parametern a från den sista ekvationen:

$$a^*_{ML,obs} = \frac{2n}{\sum_{i=1}^{n} x_i} = 2/\bar{x}.$$

Svar: $a^*_{ML,obs} = 2/\bar{x}$

b) Antag att $a > 0$. För att sätta upp funktionen $Q = Q(a)$ som ska minimeras, behöver man först beräkna väntevärdet $E(X)$. Vi har

$$\mu = E(X) = \int_0^\infty x f_X(x)\mathrm{d}x = \int_0^\infty a^2 x^2 e^{-ax}\mathrm{d}x = a\int_0^\infty x^2\big(ae^{-ax}\big)\mathrm{d}x = aE(Y^2),$$

där s.v. $Y \in Exp(a)$. Vidare, för $Y \in Exp(a)$ gäller att:

$$E(Y^2) = V(Y) + \big(E(Y)\big)^2 = |formelsamling| = \frac{1}{a^2} + \left(\frac{1}{a}\right)^2 = \frac{2}{a^2}.$$

Således,

$$\mu = E(X) = aE(Y^2) = \frac{2a}{a^2} = \frac{2}{a}.$$

Vi fortsätter med att sätta upp funktionen $Q(a)$ som ska minimeras sedan m.a.p. parametern a:

$$\begin{aligned} Q(a) &= \sum_{i=1}^n \big(x_i - \mu_i(a)\big)^2 = |\mu_i(a) = \mu = E(X) = \frac{2}{a}, \forall i| \\ &= \sum_{i=1}^n \left(x_i - \frac{2}{a}\right)^2. \end{aligned}$$

Slutligen, för att bestämma MK-skattningen $a^*_{MK,obs}$ av parametern a, dvs hitta en sådan a som minimerar funktionen $Q(a)$, deriverar vi först $Q(a)$ m.a.p. a. Sedan sätter vi upp följande ekvation och löser den för a:

$$\begin{aligned} \frac{\mathrm{d}Q(a)}{\mathrm{d}a} &= 0, \text{ som blir} \\ 2\sum_{i=1}^n \left(\left(x_i - \frac{2}{a}\right)\frac{2}{a^2}\right) &= 0, \\ \frac{4}{a^2}\sum_{i=1}^n \left(x_i - \frac{2}{a}\right) &= 0, \\ \sum_{i=1}^n \left(x_i - \frac{2}{a}\right) &= 0, \\ \sum_{i=1}^n x_i - \frac{2n}{a} &= 0. \end{aligned}$$

Således, blir MK-skattningen av parametern a från den sista ekavionen:

$$a^*_{MK,obs} = \frac{2n}{\sum_{i=1}^n x_i} = 2/\bar{x} = a^*_{ML,obs}.$$

Svar: $a^*_{MK,obs} = 2/\bar{x}$

Uppgift 16

a) Vi vet att $\sigma = 2$ och således $\bar{X} \in N(\mu, \frac{\sigma}{\sqrt{n}})$ (eftersom alla reaktionstider antas vara normalfördelade), som medför efter tranformationen att $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \in N(0,1)$.
Därför har vi:

$$
\begin{aligned}
P(\text{förkasta } H_0 | H_0 \text{ är sann}) &= 0.01 (= \alpha) \\
P(Z < -\lambda_{0.01} | \mu = 25) &= 0.01 \text{ (vänstra } \alpha\text{-svansen i } N(0,1)\text{-förd)} \\
P\Big(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < -\lambda_{0.01} | \mu = 25\Big) &= 0.01 \\
P\Big(\frac{\bar{X}-25}{2/\sqrt{n}} < -\lambda_{0.01}\Big) &= 0.01 \\
P\big(\bar{X} < 25 - \lambda_{0.01}\frac{2}{\sqrt{n}}\big) &= 0.01
\end{aligned}
$$

Vilket innebär att, baserat på det observerade medelvärdet $\bar{x}$, om

$$\bar{x} < 25 - \lambda_{0.01}\frac{2}{\sqrt{n}},$$

så förkastas $H_0 : \mu = 25$ mot $H_1 : \mu < 25$ på 1%-signifikansnivån, där $\lambda_{0.01} = 2.3263$ från Tabell 2.

Svar: Förkasta H_0 om $\bar{x} < 25 - \frac{(2.3263)2}{\sqrt{n}}$.

b) Vi söker n så att styrka $h(23) = 0.98$, eller $P(\text{förkasta } H_0 | \mu = 23) = 0.98$
Vidare, från a) och eftersom $\bar{X} \in N(\mu, \frac{\sigma}{\sqrt{n}})$, har vi:

$$
\begin{aligned}
P\big(\bar{X} < 25 - \lambda_{0.01}\frac{2}{\sqrt{n}} | \mu = 23\big) &= 0.98, \text{ som blir,} \\
P\big(\bar{X} < 25 - \lambda_{0.01}\frac{2}{\sqrt{n}} | \bar{X} \in N(23, \frac{2}{\sqrt{n}})\big) &= 0.98, \text{ vilket medför,} \\
\Phi\left(\frac{25 - \lambda_{0.01}\frac{2}{\sqrt{n}} - 23}{2/\sqrt{n}}\right) &= 0.98, \text{ och från Tabell 1,} \\
\frac{25 - \lambda_{0.01}\frac{2}{\sqrt{n}} - 23}{2/\sqrt{n}} &= 2.06.
\end{aligned}
$$

Genom att lösa den erhållna ekvationen, får vi n, dvs:

$$
\begin{aligned}
\frac{25 - \lambda_{0.01}\frac{2}{\sqrt{n}} - 23}{2/\sqrt{n}} &= 2.06, \text{ som blir,} \\
25 - 23 &= (2.06 + \lambda_{0.01})\frac{2}{\sqrt{n}}, \text{ från Tabell 2 är } \lambda_{0.01} = 2.3263, \\
2 &= (2.06 + 2.3263)\frac{2}{\sqrt{n}}, \text{ eller,} \\
\sqrt{n} &= (2.06 + 2.3263)\frac{2}{2} = 4.3863, \text{ vilket innebär att} \\
n &= 4.3863^2 \simeq 19.24
\end{aligned}
$$

Svar: Det behövs minst $n = 20$ observationer.