Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF1912/1914/1915/1916 SANNOLIKHETSTEORI OCH STATISTIK, MÅNDAG 23 OKTOBER 2023 KL 8.00–13.00.

Examinator för SF1912: Mykola Shykula, 08-790 6644.
Examinator för SF1914/1915/1916: Björn-Olof Skytt, 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 datorlaborationen behöver ej besvara uppgift 12, utan får tillgodoräkna sig denna uppgift (i svarsblanketten anges ordet Bonus). Detta gäller på den här tentan och vid omtentamen i december 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 datorlaborationen får 3 bonuspoäng på del II på ordinarie tentamenstillfället och vid omtentamen i december 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

Låt A_1 and A_2 vara två disjunkta händelser, där $A_1 \cup A_2 = \Omega$, och Ω är hela utfallsrummet. Vi vet att $P(A_1) = 0.35$. Anta nu händelsen B sådan att $P(B|A_1) = 0.75$ och $P(B \cap A_2) = 0.3$. Beräkna $P(A_2|B)$.

A: 0.4615

B: 0.5333

C: 0.2999

D: 0.5624

Uppgift 2

Låt X och Y vara två oberoende stokastiska variabler. Vi vet att $D(X) = 1.5$, $D(Y) = 2.5$ och att $D(2X + aY + 7) = 8$. Vilket värde på a är korrekt?

A: -2.96648

B: 8.8

C: 9.52

D: 3.085

Uppgift 3

Låt X vara en kontinuerlig stokastisk variabel med täthetsfunktionen

$$f_X(x) = \begin{cases} x + \frac{1}{2};\ 0 \leq x \leq 1, \\ 0; \text{annars.} \end{cases}$$

Beräkna $E(3X - 2)$.

A: 0.3333

B: 0.75

C: 1.75

D: -0.25

Uppgift 4

Antag att vi har två kontinuerliga stokastiska variabler X and Y med den simultana täthetsfunktionen.

$$f_{X,Y}(x, y) = \begin{cases} 6xy^2;\ \text{om } 0 < x < 1, 0 < y < 1, \\ 0;\ \text{annars} \end{cases}$$

Beräkna $P(\frac{1}{4} < X < \frac{1}{2})$.

A: 0.25

B: 0.375

C: 0.5625

D: 0.1875

Uppgift 5

Tiden X som en kvinna är gravid kan antas vara Normalfördelad med väntevärde μ och standardavvikelse $\sigma = 16$ dagar. Om vi vet att 67% av graviditeterna varar högst 274 dagar, vad är då förväntat värde (avrundat till hela dagar) för längden av en graviditet. Dvs. vilket värde har μ ?

A: 281

B: 258

C: 267

D: 161

Uppgift 6

Enligt ett godisföretag är 75% av deras färgade godsaker blåa hos deras mest framgångsrika märke som säljer färgade godsaker. I ett givet prov på 15 av dessa godsaker visar sig 10 vara blåa. Om vi antar att den påstådda andelen på 75% är korrekt, vad är då sannolikheten (avrundat till två decimaler) för att slumpmässigt välja 15 godsaker och få 10 eller fler som är blåa?

A: 0.85

B: 1.00

C: 0.69

D: 0.98

Uppgift 7

Låt X och Y vara två oberoende s.v. sådana att $X \in Po(\mu)$ och $Y \in Po(3\mu)$.

$$\mu^*_{obs} = \frac{2x + y}{5}$$

är en skattning av μ.
Bestäm medelfelet för denna skattning av μ om vi fått utfallen $x = 423$ och $y = 1304$.

A: 4.8

B: 11.0

C: 120.4

D: 227.5

Uppgift 8

Man har två oberoende stickprov. Stickprov 1, som består av 60 mätvärden $(x_1, x_2, \ldots, x_{60})$ kommer från en exponentialfördelning med väntevärde 3θ och stickprov 2 med 80 mätvärden $(y_1, y_2, \ldots, y_{80})$ kommer från en exponentialfördelning med väntevärde 4θ.
Data: $\bar{x} = 4.30$, $\bar{y} = 5.76$.

För att skatta θ kan man t ex använda $\theta^*_{1_{obs}} = \bar{y} - \bar{x}$ eller $\theta^*_{2_{obs}} = (\bar{x} + \bar{y})/7$.
Vilket av nedanstående påståenden är sant?

A: $\theta^*_{1_{obs}}$ är den effektivaste skattningen av θ.

B: $\theta^*_{2_{obs}}$ är den effektivaste skattningen av θ.

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 9

Efter noggrann kalibrering av en digital voltmeter fann man att mätfelen vid mätning av resistans kunde betraktas som oberoende normalfördelade stokastiska variabler med väntevärde 0 Ω och standardavvikelse 0.063 Ω. Vid upprepade mätningar av resistanserna R_1 och R_2 hos två elektriska motstånd fick man resultaten

Motstånd 1:	1.026	1.009	1.056	1.012	0.981
Motstånd 2:	0.991	1.082	1.106	1.012	1.121

Bilda ett tvåsidigt konfidensintervall som har konfidensgraden 95 % för resistansskillnaden $R_1 - R_2$ och ange övre gränsen för detta.

A: -0.0059

B: 0.0104

C: 0.0204

D: 0.0325

Uppgift 10

Man vill bygga ett stort bostadsområde med hyresrätter. Man har följande nollhypotes H_0: 10% av hyresgästerna vill ha linoleummatta i duschrummet, 40% vill ha klinkers i duschrummet, och 50% vill ha både klinkers och värmeslingor i duschrummet. 50 potentiella hyresgäster tillfrågas om hur de vill ha sina duschrum. Det visar sig att 12 av de tillfrågade potentiella hyresgästerna vill ha linoleummatta, 33 vill ha klinkers och resten vill ha både klinkers och värmeslingor i sina duschrum.
Vilket av följande påståenden är sant?

A: H_0 kan varken förkastas på risknivån 1% eller risknivån 5%

B: H_0 kan både förkastas på risknivån 1% och risknivån 5%

C: H_0 kan förkastas på risknivån 1%, men inte på risknivån 5%

D: H_0 kan förkastas på risknivån 5%, men inte på risknivån 1%

Uppgift 11

Anta att $X \in \mathrm{N}(\mu, \sigma)$ med den kända standardavvikelsen $\sigma = 1.5$ och låt H_0 vara att $\mu = 4$. Vi vill testa H_0 mot alternativet $H_1 : \mu = 6$ och förkastar H_0 till förmån för mothypotesen H_1 om vi får testvariabeln $\bar{x} \geq 5.17$ då vi gjort 9 observationer på X. Bestäm testets styrka.

A: 0.01

B: 0.71

C: 0.95

D: 0.99

Uppgift 12

Guldhalt i havsvatten bestämdes medelst en atomabsorptionspektropi-metod. Halten är mycket låg och de direkta bestämningarna osäkra. Därför använde man sig av den s k "standard addition method". Det innebär att man tillför kända halter av guld och mäter absorptionen. Absorptionen antas vara en linjär funktion av guldhalten dvs om μ är halten i rent havsvatten antas att

$$y = \beta \cdot (x + \mu) + \varepsilon,$$

eller med $\alpha = \beta \cdot \mu$:

$$y = \alpha + \beta \cdot x + \varepsilon,$$

där y är uppmätt absorption, x tillsatt mängd guld och ε mätfel som antas vara $N(0, \sigma)$.
Resultat vid undersökningen:

Tillsatt guld, ng (x)	0	10	20	30	40	50	60	70
Absorption (y)	0.257	0.314	0.364	0.413	0.468	0.528	0.574	0.634

Hjälpsummor: $\sum(x_i - \bar{x})^2 = 4200$, $\sum(y_i - \bar{y})^2 = 0.119862$, $\sum x_i y_i = 146.75$.

Skatta på ett lämpligt sätt guldkoncentrationen μ i rent havsvatten.

Ledning: Bestäm först Minsta-kvadrat-skattningarna α^*_{obs} och β^*_{obs} av linjär regressions koefficienterna α resp β.

A: Skattningen av μ är ungefär 0.048

B: Skattningen av μ är ungefär 0.48

C: Skattningen av μ är ungefär 4.8

D: Skattningen av μ är ungefär 48

Del II

Uppgift 13

Vi räknar antalet förfrågningar som tas emot på en webbserver under en timme och klassificerar dem som: X = antal förfrågningar från människor under en timme, och Y = antal förfrågningar från bots under en timme. Enligt vår information är det förväntade antalet mänskliga förfrågningar per timme 1 medan det förväntade antalet botsförfrågningar är 4.

a) Ange lämpliga probalistiska modeller för de stokastiska variablerna X och Y? (2 p)

b) Vad är sannolikhetsfunktionen för det totala antalet (från människor och från bots) förfrågningar per timme? (2 p)

c) Vad är sannolikheten för att det tagits emot 2 mänskliga förfrågningar på en viss timme givet att det tagits emot totalt 3 förfrågningar den timmen? (6 p)

Uppgift 14

Man har vid 10 olika tillfällen försökt att uppskatta antalet fortkörare genom att räkna antalet passerande bilar till och med den första fortköraren. Man erhöll följande observationsserie:

$$3 \quad 1 \quad 3 \quad 9 \quad 2 \quad 6 \quad 5 \quad 1 \quad 3 \quad 7.$$

Angiv, givetvis med motiveringar, en lämplig statistisk modell samt häled Maximum-Likelihood-skattningen av sannolikheten för att föraren av en godtyckligt vald bil skall vara en fortkörare och ange dess numeriska värde för ovanstående mätdata. Olika bilars hastighet antages vara oberoende och likafördelade. (10 p)

Uppgift 15

Livslängderna i år för en viss sorts elektroniska komponenter är oberoende och Weibullfördelade med fördelningsfunktionen

$$F_X(x) = 1 - e^{-(x/5)^c}, \quad x > 0.$$

Hundra komponenter sattes i drift samtidigt och efter 3 år fungerade fortfarande åttio av dem. Punktskatta parametern c och konstruera ett approximativt 95 % konfidensintervall för c. (10 p)

Ledning: Undersök först sannolikheten p för att en komponent håller i tre år.

Uppgift 16

X_1, X_2 och X_3 är oberoende lika fördelade stokastiska variabler med täthetsfunktionen:

$$f_{X_i}(x) = \begin{cases} \lambda e^{-\lambda x} & \text{för} \quad x \geq 0, \\ 0 & \text{annars,} \end{cases}$$

där $\lambda > 0$ är en okänd konstant, $i = 1, 2, 3$. Låt $U = \max(X_1, X_2, X_3)$ och $V = \min(X_1, X_2, X_3)$. Beräkna väntevärdet $E(U - V)$. (10 p)

Lycka till!

Avd. Matematisk statistik

KTH Matematik

LÖSNINGSFÖRSLAG TENTAMEN I SF1912/SF1914/SF1915/SF1916
SANNOLIKHETSTEORI OCH STATISTIK,
MÅNDAG 23 OKTOBER 2022 KL 8.00–13.00.

Del I, Svar

1. B

2. A

3. D

4. D

5. C

6. A

7. B

8. B

9. D

10. B

11. C

12. D

Del I, Lösningsförslag

Uppgift 1

P.g.a. Bayes sats vet vi att $P(A_2|B) = \dfrac{P(B|A_2)P(A_2)}{P(B)}$.

Eftersom $A_1 \cup A_2 = \Omega$, har vi $P(A_2) = 1 - P(A_1) = 1 - 0.35 = 0.65$.

Dessutom har vi att $P(B|A_2) = \dfrac{P(B \cap A_2)}{P(A_2)} = \dfrac{0.3}{0.65} = 0.4615$.

Nu kan vi beräkna $P(B)$ med hjälp av lagen om total sannolikhet :

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) = 0.35 \cdot 0.75 + 0.65 \cdot 0.4615 = 0.5624.$$

Slutligen fås $P(A_2|B) = \dfrac{0.4615 \cdot 0.65}{0.5624} = 0.5333$.

Uppgift 2

Vi använder oss av följande: $V(2X+aY+7) = 2^2V(X)+|a|^2V(Y)$, eftersom X and Y är oberoende.

Vi har : $V(X) = (D(X))^2 = 1.5^2 = 2.25, V(Y) = (D(Y))^2 = 2.5^2 = 6.25$ and $V(2X + aY + 7) = (D(2X + aY + 7))^2 = 8^2 = 64$.

Detta medför att $V(2X + aY + 7) = 2^2 \cdot 2.25 + |a|^2 \cdot 6.25 = 64$, vilket ger $|a|^2 = 8.8$ och $a = \pm\sqrt{8.8} = -2.966$.

Uppgift 3

Vi har följande:

$$\begin{aligned} E(3X-2) &= \int_{-\infty}^{-\infty} (3x-2) f_X(x)dx = \int_0^1 (3x-2)\left(x + \frac{1}{2}\right)dx \\ &= \int_0^1 (3x^2 - 2x + \frac{3}{2}x - 1)dx = x^3 - x^2 + \frac{3}{4}x^2 - x\Big|_0^1 \\ &= -\frac{1}{4} = -0.25. \end{aligned}$$

Uppgift 4

Vi har

$$f_{X,Y}(x,y) = \begin{cases} 6xy^2, \text{ om } 0 < x < 1, 0 < y < 1, \\ 0, \text{ annars.} \end{cases}$$

Då gäller att

$$\begin{aligned} f_X(x) &= \int_0^1 f_{X,Y}(x,y)dy = \int_0^1 6xy^2 dy \\ &= 6x \cdot \frac{y^3}{3}\Big|_0^1 = 2x. \end{aligned}$$

Då får vi följande:

$$\begin{aligned} P\left(\frac{1}{4} < X < \frac{1}{2}\right) &= \int_{1/4}^{1/2} f_X(x)dx = \int_{1/4}^{1/2} 2xdx \\ &= x^2\Big|_{1/4}^{1/2} = \frac{1}{4} - \frac{1}{16} = \frac{3}{16} = 0.1875 \end{aligned}$$

Uppgift 5

Vi vet att $X \in N(\mu, 16)$ och att $P(X \leq 274) = 0.67$. Vi gör om till N(0,1) och får $Z = \dfrac{X-\mu}{16}$.

Detta medför att

$$P(X < 274) = P\left(Z < \frac{274-\mu}{16}\right) = 0.67.$$

Å andra sidan, eftersom $Z \in N(0,1)$, så vet vi att $P(Z < z) = 0.67$ för $z = 0.44$ (Table 1).

Alltså gäller att $\dfrac{274-\mu}{16} = 0.44 \Rightarrow \mu = 274 - 0.44 \cdot 16 = 266.96 \approx 267.$

Uppgift 6

Låt X vara det antal blåa godsaker vi drar. Då gäller att $X \in Bin(n,p)$ med $n = 100$ och $p = 0.75$.

Vi vill beräkna $P(X \geq 10)$. Vi vet att om $X \in Bin(n,p)$ så gäller följande om Y är antalet icke-blåa godsaker vi drar:

$$P(X \geq 10) = P(Y \leq 15 - 10) = P(Y \leq 5), \text{ för } Y \in Bin(n, 1-p) = Bin(15, 0.25).$$

Nu kan vi använda tabell 6: $P(X \geq 10) = P(Y \leq 5) = 0.8516 \approx 0.85.$

Uppgift 7

X och Y är två oberoende s.v. sådana att $X \in Po(\mu)$ och $Y \in Po(3\mu)$.

$$\mu^*_{obs} = \frac{2x+y}{5}$$

är vår skattning av μ.
Se F.S. §9.3. Medelfelet är den skattade standardavvikelsen av skattningen.

$$V[\mu^*] = V[\frac{2X+Y}{5}] = \frac{V[2X+Y]}{25} = \textbf{ober} = \frac{V[2X]+V[Y]}{25} = \frac{4V[X]+V[Y]}{25} = \frac{4\mu+3\mu}{25} = \frac{7\mu}{25}$$

$$D[\mu^*] = \sqrt{\frac{7\mu}{25}}$$

Medelfelet för denna skattning är då

$$D^*_{obs}[\mu^*] = \sqrt{\mu^*_{obs}\frac{7}{25}} = \sqrt{\frac{2x+y}{5}\cdot\frac{7}{25}} = \sqrt{\frac{2\cdot 423+1304}{5}\cdot\frac{7}{25}} = 11.0$$

om vi fått utfallen $x = 423$ och $y = 1304$.

Uppgift 8

$$E(\theta^*_1) = E(\bar{Y}) - E(\bar{X}) = \frac{1}{80}\sum_{i=1}^{80} E(Y_i) - \frac{1}{60}\sum_{i=1}^{60} E(X_i) = \frac{1}{80}\cdot 80\cdot 4\theta - \frac{1}{60}\cdot 60\cdot 3\theta = \theta$$

visar att θ^*_1 är väntevärdesriktig.
För θ^*_2 får vi

$$E\left(\frac{\bar{X}+\bar{Y}}{7}\right) = \frac{1}{7}\left(E(\bar{X})+E(\bar{Y})\right) = \text{ (enl ovan) } = \frac{1}{7}(3\theta+4\theta) = \theta$$

vilket visar att även θ^*_2 är väntevärdesriktig.

$$V(\theta^*_1) = V(\bar{Y}-\bar{X}) = V(\bar{Y}) + V(\bar{X}) = \frac{1}{(80)^2}\sum_{1}^{80} V(Y_i) + \frac{1}{(60)^2}\sum_{1}^{60} V(X_i) =$$

$$= \text{ (Enl. formelsamlingen) } = \frac{1}{(80)^2}\cdot 80\cdot(4\theta)^2 + \frac{1}{(60)^2}\cdot 60\cdot(3\theta)^2 = \frac{7}{20}\theta^2$$

medan

$$V(\theta^*_2) = \frac{1}{7^2}V(\bar{X}) + \frac{1}{7^2}V(\bar{Y}) = \text{ (som ovan) } = \frac{(3\theta)^2}{7^2\cdot 60} + \frac{(4\theta)^2}{7^2\cdot 80} = \frac{\theta^2}{140}$$

som är mindre än $V(\theta^*_1)$ och alltså är $\theta^*_{2_{obs}}$ en effektivare skattning av θ än $\theta^*_{1_{obs}}$.

Uppgift 9

Två oberoende stickprov. Kända sigma! Vi har att $\bar{x} = 1.0168$ och $\bar{y} = 1.0624$. Ett 95 % konfidensintervall ges av

$$\bar{x} - \bar{y} \pm \lambda_{0.025}\sqrt{\frac{\sigma^2}{5} + \frac{\sigma^2}{5}} = 1.0168 - 1.0624 \pm 1.96\sqrt{\frac{0.063^2}{5} + \frac{0.063^2}{5}} = -0.0456 \pm 0.0781$$

Övre gränsen är alltså 0.0325

Uppgift 10

Situationen är den som beskrivs i avsnitt 13.10 *Test av given fördelning.* Vi har att $r = 3$ och $n = 50$. Dessutom är $x_1 = 12, x_2 = 33$ och $x_3 = 5$. Slutligen är $p_1 = \frac{1}{10}, p_2 = \frac{4}{10}$ samt $p_3 = \frac{5}{10}$. Detta ger (Alla $np_j \geq 5$)

$$Q_{\text{obs}} = \sum_{j=1}^{3} \frac{(x_j - np_j)^2}{np_j} = \frac{\left(12 - 50 \cdot \frac{1}{10}\right)^2}{50 \cdot \frac{1}{10}} + \frac{\left(33 - 50 \cdot \frac{4}{10}\right)^2}{50 \cdot \frac{4}{10}} + \frac{\left(5 - 50 \cdot \frac{5}{10}\right)^2}{50 \cdot \frac{5}{10}} = 34.25$$

5%- och 1%-kvantilerna på χ^2-fördelningen med $r - 1 = 2$ frihetsgrader är $5.99 < 34.25$ och $9.21 < 34.25$. Alltså kan vi förkasta H_0 på både 1%- och 5%-nivån.

Uppgift 11

Styrkan hos testet är $P(\text{förkasta} H_0)$ om H_1 är sann.
Dvs. vi ska ta fram $P(\bar{X} \geq 5.17)$ om $X_i \in N(6, 1.5)$

$$P(\bar{X} \geq 5.17) = [\text{Gör om till N(0,1)}] = P\left(\frac{\bar{X} - 6}{1.5/\sqrt{9}} \geq \frac{5.17 - 6}{1.5/\sqrt{9}}\right) = 1 - \Phi\left(\frac{-0.83}{0.5}\right) = 1 - (1 - \Phi(1.66)) = 0.95$$

.

Uppgift 12

Först beräknar vi: $\bar{x} = 35$, $\bar{y} = 0.444$.
Minsta-kvadrat-skattningarna
α^*_{obs} och β^*_{obs} av linjär regressions koefficienterna α resp β blir:

$$\beta^*_{obs} = \frac{\sum x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sum (x_i - \bar{x})^2} = \frac{146.75 - 8 \cdot 35 \cdot 0.444}{4200} = 0.00534,$$

och

$$\alpha^*_{obs} = \bar{y} - \beta^*_{obs}\bar{x} = 0.444 - 0.00534 \cdot 35 = 0.257$$

Alltså, har vi att den lämpliga skattningen av μ blir $\alpha^*_{obs}/\beta^*_{obs} = 0.257/0.00534 \approx 48.13$

Del II, Lösningsförslag

Uppgift 13

a) En passande probabilistisk modell för X and Y är en Poisson-fördelning , som räknar antalet händelser (slumpmässiga och oberoende) som inträffar på tidsintervallet. Eftersom vi vet att det förväntade värdet av antalet mänskliga förfrågningar är 1 att det förväntade värdet av antalet botsförfrågningar är 4 så får vi att

$$X \sim Po(1) \text{ och } Y \sim Po(4).$$

b) Metod 1: If $X \sim Po(\lambda_1)$ och $X \sim Po(\lambda_2)$ är två oberoende Poisson-fördelningar med parametrar $\lambda_1 = 1$ och $\lambda_2 = 4$, så vet vi att totala antalet förfrågningar $X + Y$ också är Poissonfördelat med parameter $\lambda_1 + \lambda_2$, dvs. $(X + Y) \sim Po(\lambda_1 + \lambda_2)$. Eftersom $Z = X + Y$ så leder det till att

$$p_Z(n) = \frac{(\lambda_1 + \lambda_2)^n}{n!} e^{-(\lambda_1+\lambda_2)} = \frac{5^n}{n!} e^{-5}.$$

Dvs $Z \sim Po(5)$

Metod 2: Vi härleder sannolikhetsfunktionen för $X+Y$ genom att använda faltningsformeln.

$$\begin{aligned}
p_Z(n) = P(X + Y = n) &= \sum_{k=0}^{n} P(X = k)P(Y = n - k) \\
&= \sum_{k=0}^{n} \frac{(\lambda_1)^k}{k!} e^{-\lambda_1} \frac{(\lambda_2)^{n-k}}{(n-k)!} e^{-\lambda_2} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{(\lambda_1)^k (\lambda_2)^{n-k}}{k!(n-k)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} (\lambda_1)^k (\lambda_2)^{n-k} \\
&= \frac{(\lambda_1 + \lambda_2)^n}{n!} e^{-(\lambda_1+\lambda_2)} = \frac{5^n}{n!} e^{-5},
\end{aligned}$$

där vi i sista likheten har använt: $(a + b)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a^k b^{n-k}$.

c) Sannolikheten för att ha haft 2 mänskliga förfrågningar under en viss timme givet att det var 3 förfrågningar totalt är:

$$\begin{aligned}
P(X = 2|X + Y = 3) &= \frac{P(X = 2 \cap X + Y = 3)}{P(X + Y = 3)} = \frac{P(X = 2 \cap Y = 1)}{P(X + Y = 3)} \\
&= \frac{P(X = 2)P(Y = 1)}{P(X + Y = 3)} = \\
&= \frac{(1)^2}{2!} e^{-1} \frac{4^1}{1!} e^{-4} \frac{3!}{e^{-5} 5^3} = \\
&= \frac{24}{250} = 0.096.
\end{aligned}$$

Uppgift 14

Observationerna $x_1, \ldots, x_{10}$ är utfall av stokastiska variabler $X_1, \ldots, X_{10}$ där

X_i = antalet bilar till och med första fortköraren vid mättillfälle nr i.

Låt p vara sannolikheten sannolikheten för att en godtyckligt vald bil skall vara en fortkörare. Då gäller att

$$p_{X_i}(k) = P[X_i = k] = (1-p)^{k-1}p$$

där $(1-p)^{k-1}$ är sannolikheten för att de första $k-1$ bilarna ej kör för fort och p är sannolikheten för att bil nr k kör för fort, dvs en ffg-fördelning ("för första gången"-fördelning). Detta eftersom de olika bilarnas hastigheter antogs vara oberoende av varandra.

Således får vi likelihoodfunktionen

$$L(p) = \prod_{i=1}^{10}(1-p)^{x_i-1}p = (1-p)^{\sum_{i=1}^{10}(x_i-1)}p^{10} = (1-p)^{\sum_{i=1}^{10} x_i - \ 10}p^{10}$$

eller

$$\ln\, L(p) = \left(\sum_{i=1}^{10} x_i - \ 10\right) \cdot \ln\,(1-p) + 10 \cdot \ln\, p$$

Derivering ger

$$\frac{d\, \ln\, L(p)}{dp} = -\frac{\sum_{i=1}^{10} x_i - \ 10}{1-p} + \frac{10}{p}\ .$$

$$\frac{d\, \ln\, L(p)}{dp} = 0 \Leftrightarrow \frac{\sum_{i=1}^{10} x_i - \ 10}{1-p} = \frac{10}{p} \Leftrightarrow 10 - 10p = p\sum_{i=1}^{10} x_i - \ 10p \Leftrightarrow p = \frac{10}{\sum_{i=1}^{10} x_i}\ .$$

Då $\ln\, L(p) \to -\infty$ då $p \to 0$ eller 1 så har vi ett maximum.

Detta ger

$$p^*_{obs_{ML}} = \frac{10}{\sum_{i=1}^{10} x_i} = \frac{10}{3+1+\cdots+7} = \frac{10}{40} = \underline{0.25}.$$

Alternativt kan man derivera Likelihoodfunktionen m.a.p p direkt.

$$\frac{dL(p)}{dp} = \frac{d\prod_{i=1}^{10}(1-p)^{x_i-1}p}{dp} = \frac{d(1-p)^{30}p^{10}}{dp} = (1-p)^{30}10p^9 - 30(1-p)^{29}p^{10} = 0$$

$$\Rightarrow (1-p)10 = 30p \Rightarrow p^*_{obs_{ML}} = \frac{10}{40} = \underline{0.25}.$$

Då $L(p) \to 0$ då $p \to 0$ eller 1 så har vi ett maximum.

Uppgift 15

Vi har $p = 1 - F(3) = e^{-(3/5)^c}$. Låt Y vara antalet komponenter som fungerar efter 3 år och $y = 80$ observationen av Y. Y är $\text{Bin}(100, p)$-fördelad och $p^*_{obs} = y/100 = 80/100 = 0.8$. Eftersom $np^*_{obs}(1 - p^*_{obs}) = 16 > 10$ fås på vanligt sätt

$$I_p = p^*_{obs} \pm 1.96\sqrt{\frac{p^*_{obs}(1 - p^*_{obs})}{100}} = 0.80 \pm 0.078 = (0.722, 0.878).$$

En naturlig skattning c^*_{obs} av c fås genom sambandet $e^{-0.6^{c^*_{obs}}} = 0.80$, dvs

$$-0.6^{c^*_{obs}} = \ln 0.80 \quad \Rightarrow \quad 0.6^{c^*_{obs}} = -\ln 0.80 = 0.223 \quad \Rightarrow \quad c^*_{obs} = \ln 0.223/\ln 0.6 = \underline{2.94}.$$

Konfidensintervallets gränser blir på analogt sätt

$$\ln(-\ln 0.722)/\ln 0.6 = 2.196 \quad \text{resp.} \quad \ln(-\ln 0.878)/\ln 0.6 = 3.992,$$

dvs $\underline{I_c = (2.196, 3.992)}$.

Uppgift 16

X_1, X_2 och X_3 är $Exp(\lambda)$-fördelade, dvs $E(X_i) = 1/\lambda, i = 1, 2, 3$.
Fördelningsfunktionen $F_U(x)$ för $U = \max(X_1, X_2, X_3)$:

$$F_U(x) = P(U \leq x) = P(\max(X_1, X_2, X_3) \leq x) = P(X_1 \leq x, X_2 \leq x, X_3 \leq x) =$$
$$= P(X_1 \leq x) \cdot P(X_2 \leq x) \cdot P(X_3 \leq x) = F_{X_1}(x)F_{X_2}(x)F_{X_3}(x) = \left(1 - e^{-\lambda x}\right)^3.$$

Därför, täthetsfunktionen $f_U(x)$ för s.v. U är:

$$f_U(x) = F_U^{'}(x) = 3\lambda e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^2.$$

Vidare, väntevärdet $E(U)$:

$$E(U) = \int_0^\infty x \cdot 3\lambda e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^2 dx = 3\lambda \int_0^\infty x\left(1 - 2e^{-\lambda x} + e^{-2\lambda x}\right)e^{-\lambda x} dx =$$
$$= 3\lambda \int_0^\infty xe^{-\lambda x} dx - 6\lambda \int_0^\infty xe^{-2\lambda x} dx + 3\lambda \int_0^\infty xe^{-3\lambda x} dx =$$
$$= \left| \int_0^\infty y\lambda e^{-\lambda y} dy = E(X_i) = \frac{1}{\lambda} \right| =$$
$$= \frac{3}{\lambda} - \frac{3}{2\lambda} + \frac{1}{3\lambda} = \frac{11}{6\lambda}.$$

Fördelningsfunktionen $F_V(x)$ för $V = \min(X_1, X_2, X_3)$:

$$F_V(x) = P(V \leq x) = P(\min(X_1, X_2, X_3) \leq x) = 1 - P(\min(X_1, X_2, X_3) \geq x) =$$
$$= 1 - P(X_1 \geq x, X_2 \geq x, X_3 \geq x) = 1 - P(X_1 \geq x)P(X_2 \geq x)P(X_3 \geq x) =$$
$$= 1 - (1 - F_{X_1}(x))(1 - F_{X_2}(x))(1 - F_{X_3}(x)) = 1 - \left(e^{-\lambda x}\right)^3 = 1 - e^{-3\lambda x}.$$

Vi ser från $F_V(x)$ att V är $Exp(3\lambda)$-fördelad stokastisk variabel och därför:

$$E(V) = \frac{1}{3\lambda}.$$

Härmed,

$$E(U - V) = E(U) - E(V) = \frac{11}{6\lambda} - \frac{1}{3\lambda} = \frac{3}{2\lambda}.$$