Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

LØSNINGSFORSLAG TIL EKSAMEN I FAG TMA4240 STATISTIKK

5.august 2004

Oppgave 1 Forurensning

X er en stokastisk variabel som angir innholdet av inhalerbart støv på en tilfeldig valgt dag.

a) Her er X er normalfordelt med forventning $\mu = 35$ og varians $\sigma^2 = 25$.

Sannsynligheten for at en måling X er over 40, $P(X > 40)$:

$$\begin{aligned} P(X > 40) &= 1 - P(X \leq 40) = 1 - P(\frac{X-\mu}{\sigma} \leq \frac{40-35}{5}) \\ &= 1 - P(Z \leq 1) = 1 - \Phi(1) = 1 - 0.8413 = \underline{\underline{0.16}} \end{aligned}$$

Sannsynligheten for at X er mellom 30 og 40, $P(30 < X < 40)$:

$$\begin{aligned} P(30 < X < 40) &= P(30 < X \leq 40) = P(X \leq 40) - P(X \leq 30) \\ &= P(\frac{X-\mu}{\sigma} \leq \frac{40-35}{5}) - P(\frac{X-\mu}{\sigma} \leq \frac{30-35}{5}) \\ &= P(Z \leq 1) - P(Z \leq -1) = \Phi(1) - \Phi(-1) = 0.8413 - 0.158 = \underline{\underline{0.68}} \end{aligned}$$

Summen av to uavhengige målinger:
Vi lar X_1 og X_2 være de to uavhengige målingene, og de er begge normalfordelt med forventning μ og varians σ^2. Da vil summen $X_1 + X_2$ også være normalfordelt med følgende forventning og varians:

$$\begin{aligned} E(X_1 + X_2) &= E(X_1) + E(X_2) = \mu + \mu = 2\mu = 35 + 35 = 70 \\ \mathrm{Var}(X_1 + X_2) &= \mathrm{Var}(X_1) + \mathrm{Var}(Y_2) = \sigma^2 + \sigma^2 = 2\sigma^2 = 25 + 25 = 50 \end{aligned}$$

Dermed blir sannsynligheten for at summen av to uavhengige målinger er over 80:

$$\begin{aligned} P(X_1 + X_2 > 80) &= P(\frac{X_1 + X_2 - 2\mu}{\sqrt{2\sigma^2}} > \frac{80-70}{\sqrt{50}}) \\ &= P(Z > 1.41) = 1 - \Phi(1.41) = 1 - 0.9207 = \underline{\underline{0.08}} \end{aligned}$$

$X_1, X_2, \ldots, X_n$ er uavhengige og identisk normalfordelte med forventning μ og varians σ^2, der både μ og σ er ukjente parametre.

Vi har målinger av $X_1, \ldots, X_5$, og får oppgitt at $\sum_{i=1}^{5} X_i = 1420$ og $\sum_{i=1}^{5}(X_i - \bar{X})^2 = 2342$.

b) En god estimator $\hat{\theta}$ er en estimator som er

- forventningsrett, dvs. $E(\hat{\theta}) = \theta$, og
- har liten varians, dvs. $\mathrm{Var}(\hat{\theta})$ er liten.

Vi liker veldig godt hvis variansen minker når antall observasjoner som estimatoren er basert på øker, og at variansen går mot 0 når antallet observasjoner går mot uendelig (konsistens).

En forventningsrett estimator $\hat{\mu}$ for μ basert på $X_1, \ldots, X_n$ er

$$\underline{\underline{\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i}}$$

Denne er forventningsrett siden

$$\mathrm{E}(\hat{\mu}) = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}n\mu = \mu$$

Variansen blir

$$\mathrm{Var}(\hat{\mu}) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \left(\frac{1}{n}\right)^2\sum_{i=1}^{n}\mathrm{Var}(X_i) = \left(\frac{1}{n}\right)^2\sum_{i=1}^{n}\sigma^2 = \left(\frac{1}{n}\right)^2 n\sigma^2 = \underline{\underline{\frac{\sigma^2}{n}}}$$

En god estimator for σ^2 er

$$\underline{\underline{S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\mu})^2}}$$

som vi setter direkte inn i variansen til $\hat{\mu}$ og får en estimator for variansen til $\hat{\mu}$:

$$\underline{\underline{\widehat{\mathrm{Var}(\hat{\mu})} = \frac{S^2}{n}}}$$

c) For å finne et 95% konfidensintervall for μ baserer vi oss på at

$$T = \frac{\hat{\mu} - \mu}{\sqrt{\frac{S^2}{n}}} = \frac{\hat{\mu} - \mu}{\frac{S}{\sqrt{n}}}$$

har en t-fordeling med $(n-1)$ frihetsgrader.

Her lar vi $\alpha = 0.05$.

$$\begin{aligned}
P(-t_{\frac{\alpha}{2},n-1} &\leq \quad T \quad \leq & t_{\frac{\alpha}{2},n-1}) = 1-\alpha \\
P(-t_{\frac{\alpha}{2},n-1} &\leq \quad \tfrac{\hat{\mu}-\mu}{\frac{S}{\sqrt{n}}} \quad \leq & t_{\frac{\alpha}{2},n-1}) = 1-\alpha \\
P(\hat{\mu} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} &\leq \quad \mu \quad \leq & \hat{\mu} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}) = 1-\alpha \\
P(\hat{\mu}_L &\leq \quad \mu \quad \leq & \hat{\mu}_U) = 1-\alpha
\end{aligned}$$

Tallsvar:

$$\begin{aligned}
\hat{\mu} &= \frac{1420}{5} = 284 \\
S &= \sqrt{\frac{2342}{4}} = 24.2 \\
t_{\frac{\alpha}{2},n-1} &= t_{\frac{0.05}{2},4} = 2.78 \\
\hat{\mu}_L &= \hat{\mu} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} = 254.0 \\
\hat{\mu}_U &= \hat{\mu} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} = 314.0
\end{aligned}$$

95% konfidensintervall for μ: $\underline{\underline{[254.0, 314.0]}}$.

d) I et hypotesetestingsproblem kan vi bestemme et forkastningsområdet slik at vi kontrollerer Type I feilen på et valgt nivå α.

$$P(\text{type I feil}) = P(\text{forkaste } H_0 | H_0 \text{ er sann}) \leq \alpha.$$

I oppgaven er det tallfestet at “Hvis $\mu \geq 300$ skal det være minst 90% sannsynlighet for å sette i verk tiltak”. Dvs. at hvis $\mu \geq 300$ skal det være maksimalt 10% sannsynlighet for å sette *ikke* sette i verk tiltak. Det betyr at hvis vi velger nullhypotesen $H_0 : \mu \geq 300$ kan vi kontrollere Type I-feilen på nivå $\alpha = 0.1$.

$$\begin{aligned}
P(\text{type I feil}) &= P(\text{forkaste } H_0 | H_0 \text{ er sann}) \\
&= P(\text{anta at } \mu < 300 \text{ og dermed ikke sette igang tiltak} | \mu \geq 300) \leq 0.1
\end{aligned}$$

Krav (i) er dermed oppfylt på grunn av nivåkravet til testen, og vi velger dermed:

$H_0 : \mu \geq 300$ mot $H_1 : \mu < 300$

med signifikansnivå $\alpha = 0.10$.

Hvis vi forkaster nullhypotesen om at $\mu \geq 300$ antar vi at $\mu < 300$ og setter derfor *ikke* i verk tiltak.

Type II feilen er gitt som "å beholde H_0 når H_0 er gal", dvs. sette igang tiltak unødvendig.

Krav (ii) svarer altså til at man i en test også ønsker sannsynligheten for Type II-feil så liten som mulig.

$$\begin{aligned} P(\text{type II feil}) &= P(\text{beholde } H_0 | H_0 \text{ er gal}) \\ &= P(\text{anta at } \mu \geq 300 \text{ og dermed sette igang tiltak} | \mu \leq 300) \end{aligned}$$

e) Vi baserer oss igjen på observatoren T fra punkt c), men velger nå

$$T_0 = \frac{\hat{\mu} - 300}{\frac{S}{\sqrt{n}}}$$

som testobservator. Vi forkaster H_0 når T_0 er liten, dvs. når T_0 er mindre enn en konstant k. Vi velger k slik at Type I-feilen kontrolleres på nivå α.

$$\begin{aligned} P(\text{forkaste } H_0 | H_0 \text{ er sann}) &\leq \alpha \\ P(T_0 \leq k | \mu \geq 300) &\leq \alpha \\ P(\frac{\hat{\mu} - 300}{\frac{S}{\sqrt{n}}} \leq k | \mu = 300) &= \alpha \\ P(\frac{\hat{\mu} - 300}{\frac{S}{\sqrt{n}}} \leq k) &= \alpha \end{aligned}$$

Når $\mu = 300$ er $T_0 = \frac{\hat{\mu}-300}{\frac{S}{\sqrt{n}}}$ t-fordelt med $n-1$ frihetsgrader og tallet k som har areal α til venstre i t-fordelingen er kvantilen $-t_{\alpha,n-1}$, dvs. $k = -t_{\alpha,n-1}$.

i) Dvs. vi forkaster H_0 når

$$\underline{\underline{T_0 = \frac{\hat{\mu} - 300}{\frac{S}{\sqrt{n}}} \leq -t_{\alpha,n-1}}}$$

ii) Alternativt kan vi løse ut forkastningsområdet over som

$$\underline{\underline{\hat{\mu} \leq 300 - t_{\alpha,n-1} \frac{S}{\sqrt{n}}}}$$

For $\alpha = 0.1$ og $n = 5$ er $-t_{0.1,4} = -1.533$. Videre har vi $T_0 = \frac{\hat{\mu}-300}{\frac{S}{\sqrt{n}}} = \frac{284-300}{\frac{24.2}{\sqrt{5}}} = -1.48$. Vi kan bruke begge måtene for å skrive opp forkastningsområdet:

i) $T_0 = \frac{\hat{\mu}-300}{\frac{S}{\sqrt{n}}} = \frac{284-300}{\frac{24.2}{\sqrt{5}}} = -1.48$. som er større enn -1.533, og dermed *ikke* gir forkastning.

ii) $\hat{\mu} = 284$ og forkastningsområdet $300 - t_{\alpha,n-1}\frac{S}{\sqrt{n}} = 300 - 1.533 \cdot \frac{24.2}{\sqrt{5}} = 283.4$. Her er $284 > 283.4$ og vi forkaster *ikke* H_0.

Konklusjonen er at det vi har observert (eller noe verre), dvs. $T_0 = -1.48$ er ganske sannsynlig (har høyere sannsynlighet enn 0.1) når H_0 er sann, og vi forkaster dermed ikke H_0. Det betyr at kommunen må sette igang tiltak.

Liten digresjon: $P-$verdien angir sannsynligheten for det vi har observert eller noe verre gitt at H_0 er sann (der verre henspeiler på den alternative hypotesen). Vi har ikke forkastet H_0 på signifikansnivå 0.01, det betyr at det vi har observert eller noe verre har større sannsynlighet enn 0.1 når H_0 er sann. Det betyr at p-verdien til testen vil være *større* enn 0.1. Regner man ut p-verdien (da må vi bruke kalkulator eller statistikk-program på datamaskinen) så er den 0.21.

f) Vi kaller den stokastiske variabelen som angir målingen på den sjette stasjonen for X_6, og lar $\hat{\mu} = \frac{1}{5}\sum_{i=1}^{5} X_i$.

Et (1-α)100% prediksjonsintervall for X_5 kan utledes ved å se på fordelingen til $(\hat{\mu} - X_6)$.

Siden $X_1, ..., X_6$ er uavhengige og normalfordelte med samme forventning μ og samme varians σ^2 vil $\hat{\mu}$ og X_6 være uavhengige (som er et viktig poeng her). Dermed vil $(\hat{\mu} - X_6)$ være normalfordelt, med

$$\begin{aligned}
\mathrm{E}(\hat{\mu} - X_6) &= \mathrm{E}(\hat{\mu}) - \mathrm{E}(X_6) = \mu - \mu = 0 \\
\mathrm{Var}(\hat{\mu} - X_6) &= \mathrm{Var}(\hat{\mu}) + \mathrm{Var}(X_6) = \frac{\sigma^2}{5} + \sigma^2 = (1 + \frac{1}{5})\sigma^2
\end{aligned}$$

Dermed vil $\frac{\hat{\mu}-X_6}{\sqrt{(1+\frac{1}{5})\sigma}}$ være standard normalfordelt, og derfor $\frac{\hat{\mu}-X_6}{\sqrt{(1+\frac{1}{5})S}}$ være t-fordelt med $n-1$ frihetsgrader. Vi kan da lage et intervall for X_6:

$$\begin{aligned}
P(-t_{\frac{\alpha}{2},n-1} \quad &\leq \quad \frac{\hat{\mu}-X_6}{\sqrt{(1+\frac{1}{5})S}} \quad \leq \quad t_{\frac{\alpha}{2},n-1}) = 1-\alpha \\
P(\hat{\mu} - t_{\frac{\alpha}{2},n-1}S\sqrt{1+\frac{1}{5}} \quad &\leq \quad X_6 \quad \leq \quad \hat{\mu} + t_{\frac{\alpha}{2},n-1}S\sqrt{1+\frac{1}{5}}) = 1-\alpha
\end{aligned}$$

Intervallet $[\hat{\mu} - t_{\frac{\alpha}{2},n-1}S\sqrt{1+\frac{1}{5}}, \hat{\mu} - t_{\frac{\alpha}{2},n-1}S\sqrt{1+\frac{1}{5}}]$ vil med 95% sannsynlighet inneholde den ukjente målingen, X_6.

Dette intervallet kaller vi et prediksjonsintervall, fordi det er en prediksjon av en uobservert stokastisk variabel. Dette i motsetning til et konfidensintervall som er et intervall for en ukjent parameter.

Prediksjonsintervallet er bredere enn konfidensintervallet. Vi ser at bredden til konfidensintervallet fra punkt c) er

$$\hat{\mu}_U - \hat{\mu}_L = 2 \cdot t_{\frac{\alpha}{2},5-1} S \sqrt{\frac{1}{5}},$$

mens predisjonsintervallet har bredde

$$2 \cdot t_{\frac{\alpha}{2},5-1} S \sqrt{(1 + \frac{1}{5})}.$$

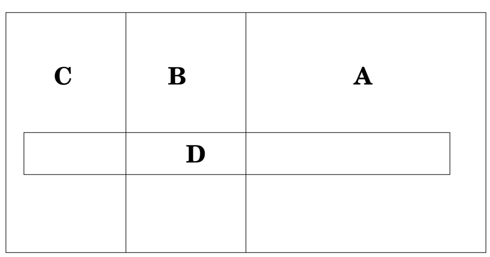
Vi er mer usikre på hvor en uobservert stokastisk variabel X_6 vil befinne seg enn vi er på hvor vi finner den ukjente parameteren μ.

Oppgave 2 Sykehjemmet

Vi ser på en tilfeldig valgt natt og definerer følgende hendelser:

$A =$ Anne er på vakt,
$B =$ Bernt er på vakt,
$C =$ Cecilie er på vakt,
$D =$ det skjer et dødsfall.
Og antar at alle dødsfall skjer naturlig.

a) Venndiagram for de fire hendelsene:



Siden det bare er Anne, Bernt og Cecilie som jobber på sykehjemmet om natten vil hendelsene A, B og C utgjøre en partisjon av utfallsrommet, og vi må ha at $P(A) + P(B) + P(C) = 1$. Dette ser vi også av venndiagrammet. Siden Bernt og Cecilie jobber like ofte må $P(B) = P(C)$. Siden Anne jobber dobbelt så ofte som hver av Bernt og Cecilie må $P(A) = 2 \cdot P(B) = 2 \cdot P(C)$. Vi utrykker alt ved $P(B)$.

$$\begin{aligned} P(A) + P(B) + P(C) &= 1 \\ 2 \cdot P(B) + P(B) + P(B) &= 1 \\ P(B) &= 0.25 \end{aligned}$$

Dermed har vi at

$$\begin{aligned} P(A) &= 0.5 \\ P(B) &= 0.25 \\ P(C) &= 0.25 \end{aligned}$$

For å regne ut $P(D)$ kan vi bruke setningen om total sannsynlighet. Vi vet at A, B, C er en partisjon av utfallsrommet.

$$\begin{aligned} P(D) &= P(D \cap A) + P(D \cap B) + P(D \cap C) \\ &= P(D|A) \cdot P(A) + P(D|B) \cdot P(B) + P(D|C) \cdot P(C) \\ &= 0.06 \cdot (0.5 + 0.25 + 0.25) = \underline{\underline{0.06}} \end{aligned}$$

Definisjonen av uavhengighet sier at C og D er to uavhengige hendelser hvis og bare hvis $P(D|C) = P(D)$, dvs. at "tilleggsinformasjon ikke endrer bildet". Vi ser fra utregningene over at $P(D|C) = P(D) = 0.06$, og C og D er dermed uavhengige hendelser.

Intuitivt vil uavhengighet av C og D følge av antagelsen om naturlig død.

b) X er en stokastisk variabel som beskriver antall av $n = 10$ naturlige dødsfall som skjer på Cecilies vakter.

Betingelser for at X er binomisk fordelt:

- Vi ser på $n = 10$ dødsfall.
- For hver dødsfall sjekker vi om Cecilie var på vakt eller ikke.
- Sannsynligheten for at Cecilie er på vakt gitt at det har skjedd et dødsfall er $P(C|D) = P(C) = 0.25$, og denne sannsynligheten er det samme for alle de n dødsfallene.
- De n dødsfallene er uavhengige siden de er naturlige (og vi antar dermed at det ikke er snakk om smittsomme sykdommer eller epidemier).

Under disse 4 betingelsene er X="antall naturlig dødsfall på Cecilies vakter" binomisk fordelt med parametere $n = 10$ og $p = 0.25$. Dermed er sannsynlighetsfordelingen til X gitt ved punktsannsynligheten $f(x)$,

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, ..., n$$

Sannsynligheten for at 7 eller flere av 10 dødsfall om natten skjer på Cecilies vakter finner vi enklest ved tabelloppslag (s 13 i formelsamlingen),

$$P(X \geq 7) \;=\; = 1 - P(X \leq 6) = 1 - 0.996 = \underline{\underline{0.004}}$$

La Y være en stokastisk variabel som angir antall sykepleiere blant 300 sykepleiere som opplever flere enn 7 dødsfall på sine vakter av totalt 10 dødsfall. Y vil dermed være binomisk fordelt med $n = 300$ og $p = 0.004$.

Sannsynligheten for at minst en av de 300 sykepleierne opplever at 7 eller flere av 10 naturlige dødsfall skjer på sine vakter er gitt som $P(Y \geq 1)$.

$$\begin{aligned} P(Y \geq 1) &= 1 - P(Y = 0) = 1 - \binom{300}{0} 0.004^0 (1 - 0.004)^{300-0} \\ &= 1 - 0.996^{300} = 1 - 0.3 = \underline{\underline{0.7}} \end{aligned}$$

Selv om det er lite sannsynlig (bare 4 promille) at det skjer 7 av 10 naturlige dødsfall på Cecilies vakter, er det svært sannsynlig (70 prosent) at minst 7 av 10 dødsfall kan skje på vakten til en av sykepleierne i Norge som jobber i samme stillingstype som Cecilie. Disse observasjonene styrker ikke mistanken mot Cecilie.

Analogi: Hver uke er det (som regel) noen som får 7 rette i Lotto, selv om dette har en forsvinnende lav sannsynlighet for hver Lotto-spiller.

Oppgave 3 Alpinulykker

Poisson-fordeling med forventningsverdi $\mu = \lambda t$, der λ skadefrekvens pr skidag og t er eksponeringstid i antall skidager.

a) I dette punktet er det kjent at $\lambda = 1/1000$ for “Alpinfjellet”.

Sannsynligheten for at det skjer akkurat én ulykke i løpet av $t = 2000$ skidager:

$$P(X = 1) \;=\; \frac{(0.001 \cdot 2000)^1}{1!} \exp(-0.001 \cdot 2000) = 2 \cdot \exp(-2) = \underline{\underline{0.27}}$$

Sannsynligheten for at du utsettes for en eller flere ulykker ved opphold 10 skidager i “Alpinfjellet”:

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - \frac{(0.001 \cdot 10)^0}{0!} \exp(-0.001 \cdot 10) = 1 - \exp(-0.01) = \underline{\underline{0.01}}. \end{aligned}$$

Sannsynligheten for minst en ulykke ved opphold i t skidager er:

$$P(X \geq 1) \;=\; 1 - P(X = 0) = 1 - \exp(-0.001 \cdot t)$$

Vi skal finne t slik at denne sannsynligheten blir større enn 0.1.

$$\begin{aligned}
P(X \geq 1) &> 0.1 \\
1 - \exp(-0.001 \cdot t) &> 0.1 \\
\exp(-0.001 \cdot t) &< 0.9 \\
-0.001 \cdot t &< \ln 0.9 \\
t &> -1000 \cdot \ln 0.9 = 105.4
\end{aligned}$$

Du må tilbringe minst $\underline{\underline{106}}$ skidager i "Alpinfjellet" for at din sannsynlighet for minst en ulykke skal bli større enn 0.1.

b) Vi har observasjoner av samhørende verdier av antall ulykker, X_i, og eksponering i antall skidager, t_i ($i = 1, ..., n$), for n tilfeldig valgte dager anlegget var åpent.

Vi finner sannsynlighetsmaksimeringsestimatoren (SME) for λ basert på de n uavhengige observasjonsparene $(X_1, t_1), (X_2, t_2), \ldots, (X_n, t_n)$. Vi ser først på rimelighetsfunksjonen, og siden observasjonsparene er uavhengige finner vi den ved å multiplisere sammen marginalsannsynlighetene. Vi innfører $f_i(x_i; \lambda)$:

$$f_i(x_i; \lambda) = \frac{(\lambda t_i)^{x_i}}{x_i!} \exp(-\lambda t_i)$$

Rimelighetsfunksjonen er gitt ved:

$$\begin{aligned}
L(\lambda) &= L(\lambda; x_1, ..., x_n) = f(x_1, ..., x_n; \lambda) \\
&= f_1(x_1; \lambda) \cdots f_n(x_n; \lambda) \\
&= \prod_{i=1}^{n} \frac{(\lambda t_i)^{x_i}}{x_i!} \exp(-\lambda t_i)
\end{aligned}$$

Tar logaritmen:

$$\begin{aligned}
l(\lambda; x_1, ..., x_n) &= \ln\left[L(\lambda)\right] \\
&= \ln\left(\prod_{i=1}^{n} \frac{1}{x_i!}\right) + \sum_{i=1}^{n} x_i \ln(\lambda t_i) - \lambda \sum_{i=1}^{n} t_i
\end{aligned}$$

Finn maksimumspunkt ved å derivere ln-rimelighetsfunksjonen og sette lik 0.

$$\frac{\partial l}{\partial \lambda} = 0 + \sum_{i=1}^{n} (x_i \frac{1}{\lambda t_i} \cdot t_i) - \sum_{i=1}^{n} t_i = \frac{1}{\lambda} \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} t_i$$

Settes dette uttrykket lik 0, får vi løsningen

$$\widehat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} t_i}$$

SME for λ blir da $\underline{\underline{\hat{\lambda} = \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} t_i}}}$.

Er $\hat{\lambda}$ forventningsrett:

$$
\begin{aligned}
\mathrm{E}[\hat{\lambda}] &= \mathrm{E}\left[\frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n t_i}\right] \quad = \quad \frac{1}{\sum_{i=1}^n t_i}\mathrm{E}(\sum_{i=1}^n X_i) \\
&= \frac{1}{\sum_{i=1}^n t_i}\sum_{i=1}^n \mathrm{E}(X_i) = \frac{1}{\sum_{i=1}^n t_i}\sum_{i=1}^n \lambda t_i \\
&= \underline{\underline{\lambda}}.
\end{aligned}
$$

Ja, estimatoren er forventningsrett.