

LØSNING: Oppgavesett nr. 6

“MAT110 Statistikk I”, 2019

Oppgave 1: (estimering - logistikk , McDonald's)

a) Forventningen av estimatoren $\hat{\lambda}$:

$$\underline{\underline{E[\hat{\lambda}]} = E[\bar{X}] = E\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n E[X_i] = \frac{n}{n} \lambda = \underline{\underline{\lambda}}}} \quad (1)$$

altså $\hat{\lambda}$ er forventningsrett.

b) Punktestimat for λ : (realisering - AIT-tid)

$$\underline{\underline{\hat{\lambda}(x_1, x_2, \dots, x_n) = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \underline{\underline{0.439 \text{ minutter}}}}} \quad (2)$$

hvor vi bruk er at $\bar{x} = 0.439$ minutter fra oppgave øving 5.

Punktestimat for μ : (realisering - forventet bearbeidingstid)

$$\underline{\underline{\hat{\mu}(y_1, y_2, \dots, y_n)}} = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \underline{\underline{6.300 \text{ minutter}}} \quad (3)$$

hvor vi bruk er at $\bar{y} = 6.300$ minutter fra oppgave øving 5.

Punktestimat for σ : (realisering - standardavvik av bearbeidingstid)

$$\underline{\underline{\hat{\sigma}(y_1, y_2, \dots, y_n)}} = s_y = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2} = \underline{\underline{1.993 \text{ minutter}}} \quad (4)$$

hvor vi bruk er at $s_y = 1.993$ minutter fra oppgave øving 5.

c) Asymptotisk $(1 - \alpha)$ 100 % konfidensintervall for

$$\hat{\mu} = \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (5)$$

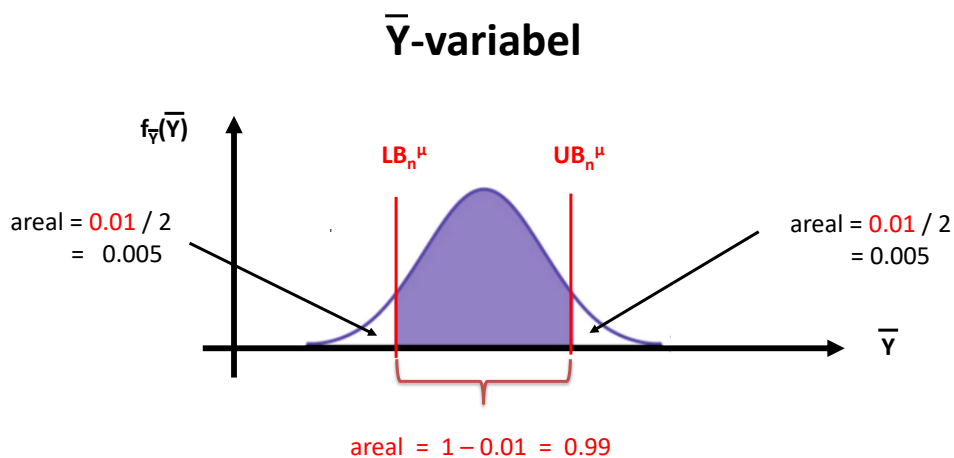
Fra oppgaven vet vi at:

$$Y_1, Y_2, Y_3, \dots, Y_n \stackrel{\text{i.i.d.}}{\sim} Y \sim N[\mu, \sigma] \quad (6)$$

som blant annet betyr at $E[Y] = E[Y_i] = \mu$ og at $\sigma^2[Y] = \sigma^2[Y_i] = \sigma^2$.

Definisjonen av et [asymptotisk](#) konfidensintervall med $\theta = \mu$:

$$\lim_{n \rightarrow \infty} P_{\mu}(LB_n^{\mu} \leq \mu \leq UB_n^{\mu}) = 1 - \alpha \quad (7)$$



Figur 1: 99 %-konfidensintervall for μ .

Alle Y_i i lign.(6) oppfyller **i.i.d.** (med $Y_i \sim N[\mu, \sigma]$).
Sum normalfordelinger er fortsatt normalfordelt:

$$\bar{Y} \sim N[E[\bar{Y}], \sigma[\bar{Y}]] \quad (8)$$

Standardiserer: ¹

$$\underline{Z} = \frac{\bar{Y} - \textcolor{red}{E}[\bar{Y}]}{\textcolor{blue}{\sigma}[\bar{Y}]} = \frac{\bar{Y} - \textcolor{red}{\mu}}{\underline{\frac{\sigma}{\sqrt{n}}}} \stackrel{n \rightarrow \infty}{\sim} N[0, 1] \quad (12)$$

¹Estimatoren $\hat{\mu}$ er forventningsrett, dvs.:

$$E[\hat{\mu}] = \textcolor{red}{E}[\bar{Y}] = \textcolor{red}{\mu} \quad (9)$$

altså estimatoren $\hat{\mu}$ er forventningsrett. Variansen til $\hat{\mu}$ er:

$$Var[\hat{\mu}] = Var[\bar{Y}] = \frac{\sigma^2}{n} \quad (10)$$

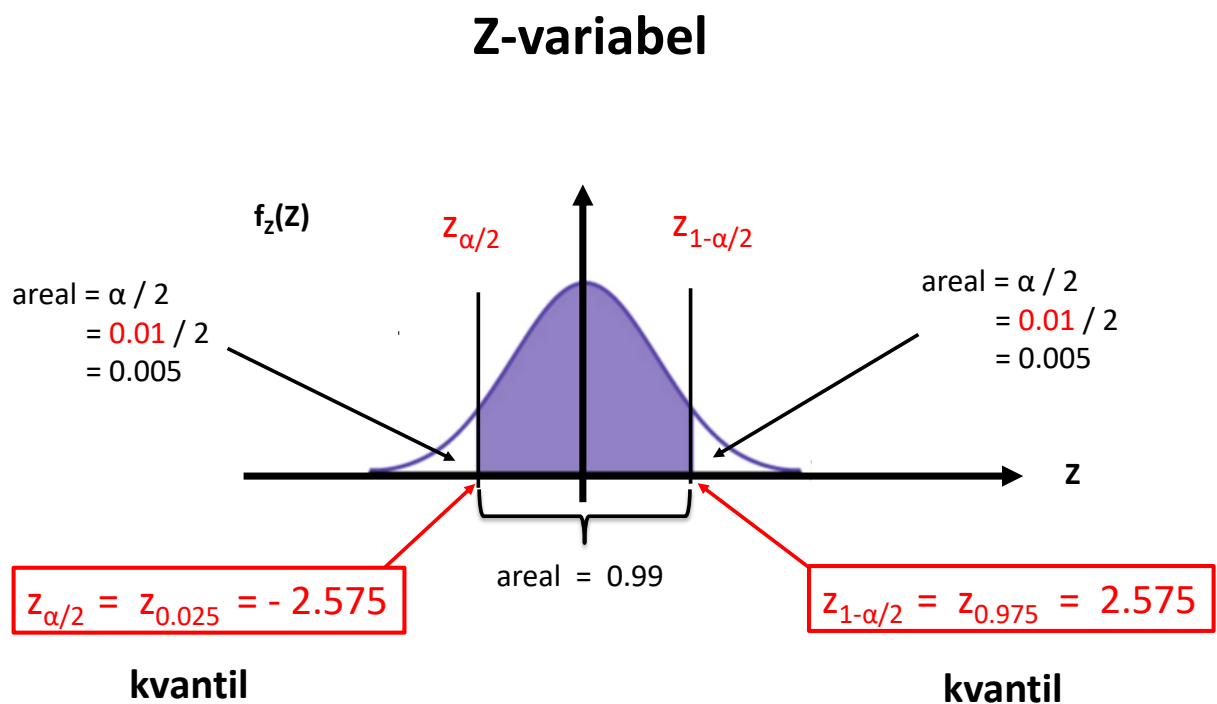
Generelt er $\sigma[Y] = \sqrt{Var[Y]}$, dermed:

$$\sigma[\hat{\mu}] = \textcolor{blue}{\sigma}[\bar{Y}] = \frac{\sigma}{\sqrt{n}} \quad (11)$$

Kvantilene når $\alpha = 0.001$: (se figur 2) ²

$$z_{\alpha/2} = z_{0.005} = -2.575 \quad , \quad z_{1-\alpha/2} = z_{0.995} = 2.575 \quad (13)$$

via omvendt tabeloppslag for normalfordelingen fra kompendiet.



Figur 2: Kvantil.

²Absoluttverdiene til $z_{\alpha/2}$ og $z_{1-\alpha/2}$ er like fordi normalfordelingen er symmetrisk.

Fra definisjonen i lign.(7):

$$\lim_{n \rightarrow \infty} P_{\mu}(LB_n^{\mu} \leq \mu \leq UB_n^{\mu}) = 1 - \alpha \quad (14)$$

Siden estimatoren $\hat{\mu}$ er normalfordelt, se lign.(8), så standardiserer vi:

$$\lim_{n \rightarrow \infty} P(z_{\alpha/2} \leq Z \leq z_{1-\alpha/2}) = 1 - \alpha \quad (15)$$

hvor Z er gitt ved lign.(79):

$$\lim_{n \rightarrow \infty} P\left(z_{\alpha/2} \leq \frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{1-\alpha/2}\right) = 1 - \alpha \quad (16)$$

$\Updownarrow$ (algebra)

$$\lim_{n \rightarrow \infty} P\left(\underbrace{\bar{Y} - \frac{z_{1-\alpha/2}}{\sqrt{n}} \sigma}_{= LB_n^{\mu}} \leq \mu \leq \underbrace{\bar{Y} - \frac{z_{\alpha/2}}{\sqrt{n}} \sigma}_{= UB_n^{\mu}}\right) = 1 - \alpha \quad (17)$$

og hvor vi definerer nedre og øvre grenser:

$$LB_n^{\mu} = \bar{Y} - \frac{z_{1-\alpha/2}}{\sqrt{n}} \sigma \quad (18)$$

$$UB_n^{\mu} = \bar{Y} - \frac{z_{\alpha/2}}{\sqrt{n}} \sigma \quad (19)$$

Med $\hat{\sigma} = S_y$ innsatt for σ i lign.(18) og (19):

$$LB_n^{\mu} = \bar{Y} - \frac{z_{1-\alpha/2}}{\sqrt{n}} S_y \quad (20)$$

$$UB_n^{\mu} = \bar{Y} - \frac{z_{\alpha/2}}{\sqrt{n}} S_y \quad (21)$$

Intervall

$$\underline{\underline{[LB_n^\mu, UB_n^\mu] = \left[\bar{Y} - \frac{z_{1-\alpha/2}}{\sqrt{n}} S_y, \bar{Y} - \frac{z_{\alpha/2}}{\sqrt{n}} S_y \right]}} \quad (22)$$

er dermed et *asymptotisk* $(1 - \alpha)$ 100 %-konfidensintervall for μ .

Fra løsningen av øving 5 har vi at:

$$\bar{y} = 6.300 \text{ minutter} \quad (23)$$

$$s_y = 1.993 \text{ minutter} \quad (24)$$

Realisering: (tall)

$$\underline{LB_n^\mu(y_1, y_2, \dots, y_n)} = \bar{y} - \frac{z_{1-\alpha/2}}{\sqrt{n}} s_y \quad (25)$$

$$= \left(6.300 - \frac{2.575}{\sqrt{100}} \cdot 1.993 \right) \text{ minutter} \quad (26)$$

$$= \underline{5.7868 \text{ minutter}} \quad (27)$$

$$\underline{UB_n^\mu(y_1, y_2, \dots, y_n)} = \bar{y} - \frac{z_{\alpha/2}}{\sqrt{n}} s_y \quad (28)$$

$$= \left(6.300 - \frac{(-2.575)}{\sqrt{100}} \cdot 1.9393 \right) \text{ minutter} \quad (29)$$

$$= \underline{6.8132 \text{ minutter}} \quad (30)$$

som gir realiseringen

$$\underline{\underline{[LB_n^\mu, UB_n^\mu]}} = [5.7868, 6.8132] \text{ minutter} \quad (31)$$

av det asymptotiske 99 %-konfidensintervallet for μ .

d) Konfidensintervallet finnes på samme måte som i oppgave 1c, se lign.(22): ³

$$\underline{\underline{[LB_n^\mu, UB_n^\mu] = \left[\bar{Y} - \frac{q_{1-\alpha/2}}{\sqrt{n}} S_y, \bar{Y} - \frac{q_{\alpha/2}}{\sqrt{n}} S_y \right]}} \quad (32)$$

hvor $q_{\alpha/2}$ og $q_{1-\alpha/2}$ er kvantilene til til Students t -fordelingen.

Kvantilene når $\alpha = 0.001$: (jfr. figur 2) ⁴

$$q_{\alpha/2} = q_{0.0005} = -2.626 \quad , \quad q_{1-\alpha/2} = z_{0.9995} = 2.626 \quad (33)$$

siden $n = 100$.

Fra løsningen av øving 5 har vi at:

$$\bar{y} = 6.300 \text{ minutter} \quad (34)$$

$$s_y = 1.993 \text{ minutter} \quad (35)$$

³Finner lign.(32) i kompendiet.

⁴Absoluttverdiene til $q_{\alpha/2}$ og $q_{1-\alpha/2}$ er like fordi Student's t -fordelingen, i likhet med normalfordelingen, er symmetrisk.

Realisering: (tall)

$$\underline{LB_n^\mu(y_1, y_2, \dots, y_n)} = \bar{y} - \frac{q_{1-\alpha/2}}{\sqrt{n}} s_y \quad (36)$$

$$= \left(6.300 - \frac{2.626}{\sqrt{100}} \cdot 1.993 \right) \text{ minutter} \quad (37)$$

$$= \underline{5.7766 \text{ minutter}} \quad (38)$$

$$\underline{UB_n^\mu(y_1, y_2, \dots, y_n)} = \bar{y} - \frac{q_{\alpha/2}}{\sqrt{n}} s_y \quad (39)$$

$$= \left(6.300 - \frac{(-2.626)}{\sqrt{100}} \cdot 1.9393 \right) \text{ minutter} \quad (40)$$

$$= \underline{6.8234 \text{ minutter}} \quad (41)$$

som gir realiseringen

$$\underline{\underline{[LB_n^\mu, UB_n^\mu]}} = [5.7766, 6.8234] \text{ minutter} \quad (42)$$

av det asymptotiske 99 %-konfidensintervallet for μ .

Siden Student's t -fordelingen har tykkere haler så er kvantilene større. Dermed blir også konfidensintervallet større.

- e) Definisjonen av et asymptotisk konfidensintervall er: (se kompendium)

$$\lim_{n \rightarrow \infty} P_{\mu}(LB_n^{\sigma} \leq \sigma \leq UB_n^{\sigma}) = 1 - \alpha \quad (43)$$

Siden Y_i -variablene er *i.i.d.* så er gjennomsnittet $\bar{Y}$ normalfordelt.

Ved å standardisere og gjøre tilsvarende som i forrige eksempel så får man:

$$[LB_n^{\sigma}, UB_n^{\sigma}] = \left[\sqrt{\frac{n-1}{(n-1) + z_{1-\alpha/2} \sqrt{2(n-1)}}} S_y, \sqrt{\frac{n-1}{(n-1) + z_{\alpha/2} \sqrt{2(n-1)}}} S_y \right] \quad (44)$$

Kvantilene når $\alpha = 0.001$: (se figur 2) ⁵

$$z_{\alpha/2} = z_{0.005} = -2.575, \quad z_{1-\alpha/2} = z_{0.995} = 2.575 \quad (45)$$

via omvendt tabelloppslag fra kompendiet.

Fra løsningen av øving 5 har vi at:

$$s_y = 1.993 \text{ minutter} \quad (46)$$

⁵Absoluttverdiene til $z_{\alpha/2}$ og $z_{1-\alpha/2}$ er like fordi N -fordelingen er symmetrisk.

Realisering: (tall)

$$\underline{LB_n^\sigma(y_1, y_2, \dots, y_n)} = \sqrt{\frac{n-1}{(n-1) + \textcolor{red}{z}_{1-\alpha/2} \sqrt{2(n-1)}}} s_y \quad (47)$$

$$= \sqrt{\frac{100-1}{(100-1) + \textcolor{red}{2.575} \sqrt{2(100-1)}}} \cdot 1.993 \quad (48)$$

$$= \underline{1.705 \text{ minutter}} \quad (49)$$

$$\underline{UB_n^\sigma(y_1, y_2, \dots, y_n)} = \sqrt{\frac{n-1}{(n-1) + \textcolor{red}{z}_{\alpha/2} \sqrt{2(n-1)}}} s_y \quad (50)$$

$$= \sqrt{\frac{100-1}{(100-1) - \textcolor{red}{2.575} \sqrt{2(100-1)}}} \cdot 1.993 \quad (51)$$

$$= \underline{2.503 \text{ minutter}} \quad (52)$$

som gir realiseringen

$$\underline{\underline{[LB_n^\sigma, UB_n^\sigma]}} = [1.705, 2.503] \text{ minutter} \quad (53)$$

av det asymptotiske 99 %-konfidensintervallet for standardavviket σ til bearbeidsingtiden Y .

- f) Siden $Y_1, Y_2, \dots, Y_n$ er *i.i.d.* normalfordelte stokastiske variabler, dvs. $Y_i \sim N[\mu, \sigma]$ så er den stokastiske variabelen

$$\frac{(n-1)S_y^2}{\sigma^2} \sim \chi_{n-1}^2 \quad (54)$$

χ_{n-1}^2 -fordelt med $n-1$ frihetsgrader.

Vi bruker χ_{n-1}^2 -fordelingens α kvantiler, $\chi_{\alpha/2}$ og $\chi_{1-\alpha/2}$, for å konstruere et eksakt 99 %-konfidensintervall for σ :

$$P\left(\chi_{\alpha/2} \leq \frac{(n-1)S_y^2}{\sigma^2} \leq \chi_{1-\alpha/2}\right) = 1 - \alpha \quad (55)$$

$\Updownarrow$ (algebra)

$$P\left(\underbrace{\sqrt{\frac{n-1}{\chi_{1-\alpha/2}}} S_y}_{= LB_n^\sigma} \leq \sigma \leq \underbrace{\sqrt{\frac{n-1}{\chi_{\alpha/2}}} S_y}_{= UB_n^\sigma}\right) = 1 - \alpha \quad (56)$$

hvor vi definerer dermed nedre og øvre grenser:

$$LB_n^\sigma = \sqrt{\frac{n-1}{\chi_{1-\alpha/2}}} S_y \quad (57)$$

$$UB_n^\sigma = \sqrt{\frac{n-1}{\chi_{\alpha/2}}} S_y \quad (58)$$

med

$$\chi_{\alpha/2} = \text{nedre kvantil til } \chi_{n-1}^2\text{-fordelingen} \quad (59)$$

$$\chi_{1-\alpha/2} = \text{øvre kvantil til } \chi_{n-1}^2\text{-fordelingen} \quad (60)$$

Intervallet

$$\underline{\underline{[LB_n^\sigma, UB_n^\sigma] = \left[\sqrt{\frac{n-1}{\chi_{1-\alpha/2}}} S_y, \sqrt{\frac{n-1}{\chi_{\alpha/2}}} S_y \right]}} \quad (61)$$

er dermed et eksakt $(1 - \alpha)$ 100 %-konfidensintervall for σ for effekten Y .

Med $\alpha = 0.01$ og $n = 100$ for bearbeidsingstiden Y så finner vi f.eks. ved å google på nettet: ⁶ ⁷

$$\chi_{\alpha/2} = \chi_{0.005} = 66.5 \quad (62)$$

$$\chi_{1-\alpha/2} = \chi_{0.995} = 139.0 \quad (63)$$

Fra løsningen av øving 5:

$$s_y = 1.993 \text{ minutter} \quad (64)$$

Realisering: (tall)

$$\underline{LB_n^\sigma(y_1, y_2, \dots, y_n)} = \sqrt{\frac{n-1}{\chi_{1-\alpha/2}}} s_y \quad (65)$$

$$= \sqrt{\frac{100-1}{139.0}} \cdot 1.993 \quad (66)$$

$$= \underline{1.6820} \quad (67)$$

$$\underline{UB_n^\sigma(y_1, y_2, \dots, y_n)} = \sqrt{\frac{n-1}{\chi_{\alpha/2}}} s_y \quad (68)$$

$$= \sqrt{\frac{100-1}{66.5}} \cdot 1.993 \quad (69)$$

$$= \underline{2.4317} \quad (70)$$

⁶Kvantilene $\chi_{\alpha/2}$ og $\chi_{1-\alpha/2}$ er forskjellige fordi χ -fordelingen ikke er symmetrisk, se f.eks. figur i kompendiet eller google på internett.

⁷Ved googling på internett så finner man mange nettsteder hvor man finner de aktuelle kvantilene, f.eks. <https://keisan.casio.com/exec/system/1180573197>. Husk at $n = 100$.

som gir realiseringen

$$\underline{\underline{[LB_n^\sigma, UB_n^\sigma]}} = [1.6820, 2.4317] \quad (71)$$

av det eksakte 99 %-konfidensintervallet for standardavviket σ til bearbeidingstiden Y .

g) Asymptotisk $(1 - \alpha)$ 100 % konfidensintervall for

$$\hat{\lambda} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (72)$$

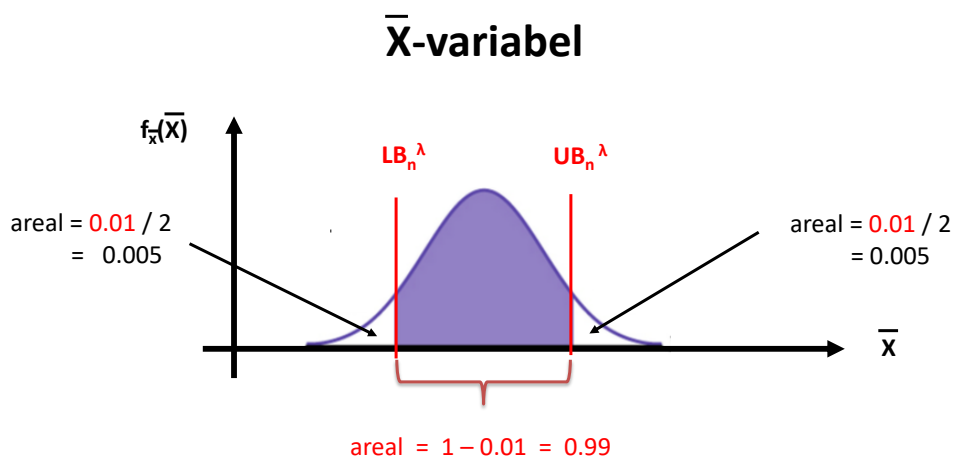
Fra oppgaven vet vi at:

$$X_1, X_2, X_3, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} X \sim \text{Exp}[\lambda] \quad (73)$$

I fotnoten i oppgaveteksten var det oppgitt at $E[X] = E[X_i] = \lambda$ og at $\sigma^2[X] = \sigma^2[X_i] = \lambda^2$.

Definisjonen av et [asymptotisk](#) konfidensintervall med $\theta = \lambda$:

$$\lim_{n \rightarrow \infty} P_{\lambda}(LB_n^{\lambda} \leq \lambda \leq UB_n^{\lambda}) = 1 - \alpha \quad (74)$$



Figur 3: Asymptotisk 99 %-konfidensintervall for λ .

Siden alle X_i i lign.(73) oppfyller **i.i.d.** (med $X_i \sim \text{Exp}[\lambda]$), får vi fra sentralgrenseteoremet at

$$\overline{X} \stackrel{n \rightarrow \infty}{\sim} N[E[\overline{X}], \sigma[\overline{X}]] \quad (75)$$

Standardiserer: ⁸

$$\underline{Z} = \frac{\overline{X} - \textcolor{red}{E}[\overline{X}]}{\sigma[\overline{X}]} = \frac{\overline{X} - \textcolor{red}{\lambda}}{\underline{\frac{\lambda}{\sqrt{n}}}} \stackrel{n \rightarrow \infty}{\sim} N[0, 1] \quad (79)$$

⁸Estimatoren $\hat{\lambda}$ er forventningsrett, dvs.:

$$E[\hat{\lambda}] = \textcolor{red}{E}[\overline{X}] = \textcolor{red}{\lambda} \quad (76)$$

Variansen til $\hat{\lambda}$ er:

$$\text{Var}[\hat{\lambda}] = \text{Var}[\overline{X}] = \frac{\lambda^2}{n} \quad (77)$$

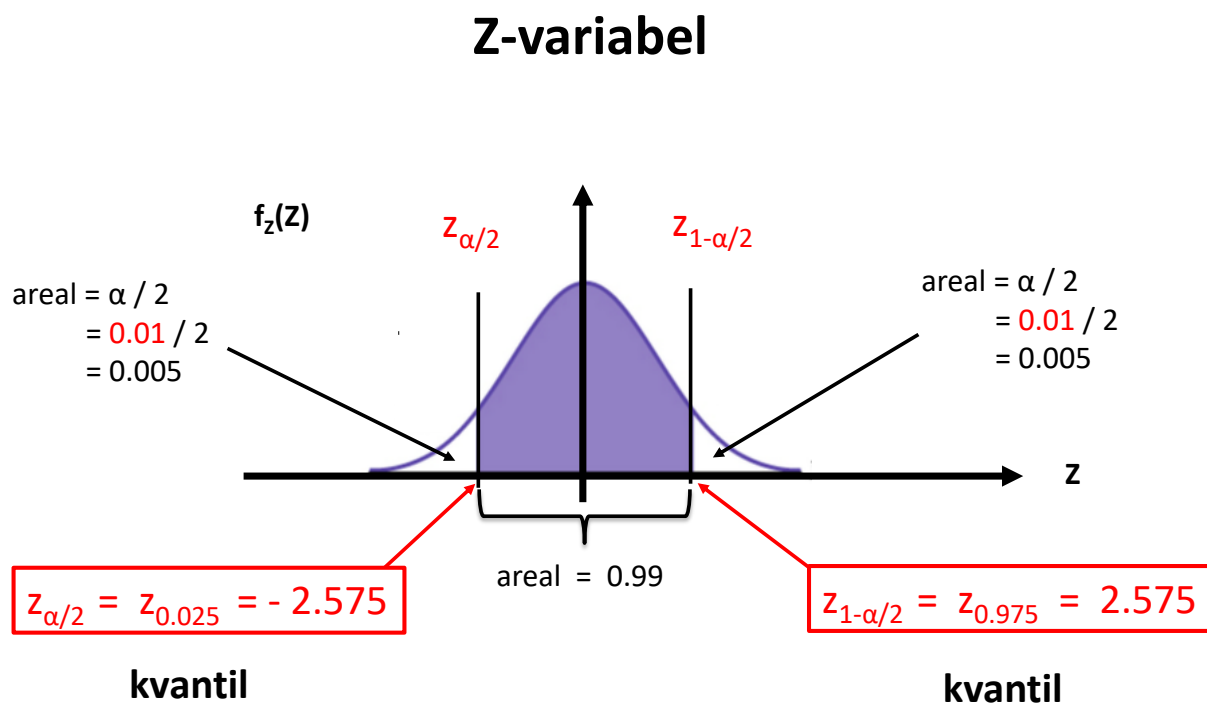
Generelt er $\sigma[X] = \sqrt{\text{Var}[X]}$, dermed:

$$\sigma[\hat{\lambda}] = \sigma[\overline{X}] = \underline{\frac{\lambda}{\sqrt{n}}} \quad (78)$$

Kvantilene når $\alpha = 0.001$: (se figur 4) ⁹

$$z_{\alpha/2} = z_{0.005} = -2.575 \quad , \quad z_{1-\alpha/2} = z_{0.995} = 2.575 \quad (80)$$

via omvendt tabeloppslag for normalfordelingen fra kompendiet.



Figur 4: Kvantil.

⁹Absoluttverdiene til $z_{\alpha/2}$ og $z_{1-\alpha/2}$ er like fordi normalfordelingen er symmetrisk.

Fra definisjonen i lign.(74):

$$\lim_{n \rightarrow \infty} P_{\lambda}(LB_n^{\lambda} \leq \lambda \leq UB_n^{\lambda}) = 1 - \alpha \quad (81)$$

Siden estimatoren $\hat{\lambda}$ er normalfordelt, se lign.(75), så standardiserer vi:

$$\lim_{n \rightarrow \infty} P(z_{\alpha/2} \leq Z \leq z_{1-\alpha/2}) = 1 - \alpha \quad (82)$$

hvor Z er gitt ved lign.(79):

$$\lim_{n \rightarrow \infty} P\left(z_{\alpha/2} \leq \frac{\bar{X} - \lambda}{\frac{\lambda}{\sqrt{n}}} \leq z_{1-\alpha/2}\right) = 1 - \alpha \quad (83)$$

$\Updownarrow$ (algebra)

$$\lim_{n \rightarrow \infty} P\left(\underbrace{\bar{X} - \frac{z_{1-\alpha/2}}{\sqrt{n}} \lambda}_{= LB_n^{\lambda}} \leq \lambda \leq \underbrace{\bar{X} - \frac{z_{\alpha/2}}{\sqrt{n}} \lambda}_{= UB_n^{\lambda}}\right) = 1 - \alpha \quad (84)$$

og hvor vi definerer nedre og øvre grenser:

$$LB_n^{\lambda} = \bar{X} - \frac{z_{1-\alpha/2}}{\sqrt{n}} \lambda \quad (85)$$

$$UB_n^{\lambda} = \bar{X} - \frac{z_{\alpha/2}}{\sqrt{n}} \lambda \quad (86)$$

Med $\hat{\lambda} = \bar{X}$ innsatt for λ i lign.(85) og (86):

$$LB_n^{\lambda} = \bar{X} - \frac{z_{1-\alpha/2}}{\sqrt{n}} \bar{X} \quad (87)$$

$$UB_n^{\lambda} = \bar{X} - \frac{z_{\alpha/2}}{\sqrt{n}} \bar{X} \quad (88)$$

Intervall

$$\underline{\underline{[LB_n^\lambda, UB_n^\lambda] = \left[\overline{X} - \frac{z_{1-\alpha/2}}{\sqrt{n}} \overline{X}, \overline{X} - \frac{z_{\alpha/2}}{\sqrt{n}} \overline{X} \right]}} \quad (89)$$

er dermed et *asymptotisk* $(1 - \alpha)$ 100 %-konfidensintervall for λ .

Fra løsningen av øving 5 har vi at:

$$\bar{x} = 0.439 \text{ minutter} \quad (90)$$

Realisering: (tall)

$$\underline{LB_n^\lambda(x_1, x_2, \dots, x_n)} = \bar{x} - \frac{z_{1-\alpha/2}}{\sqrt{n}} \bar{x} \quad (91)$$

$$= \left(0.439 - \frac{2.575}{\sqrt{100}} \cdot 0.439 \right) \text{ minutter} \quad (92)$$

$$= \underline{0.326 \text{ minutter}} \quad (93)$$

$$\underline{UB_n^\lambda(x_1, x_2, \dots, x_n)} = \bar{x} - \frac{z_{\alpha/2}}{\sqrt{n}} \bar{x} \quad (94)$$

$$= \left(0.439 - \frac{(-2.575)}{\sqrt{100}} \cdot 0.439 \right) \text{ minutter} \quad (95)$$

$$= \underline{0.552 \text{ minutter}} \quad (96)$$

som gir realiseringen

$$\underline{\underline{[LB_n^\lambda, UB_n^\lambda]}} = [0.326, 0.552] \text{ minutter} \quad (97)$$

av det asymptotiske 99 %-konfidensintervallet for λ .

■