ISBN 9788770668279

Kapitel 2

Øvelse 2.2

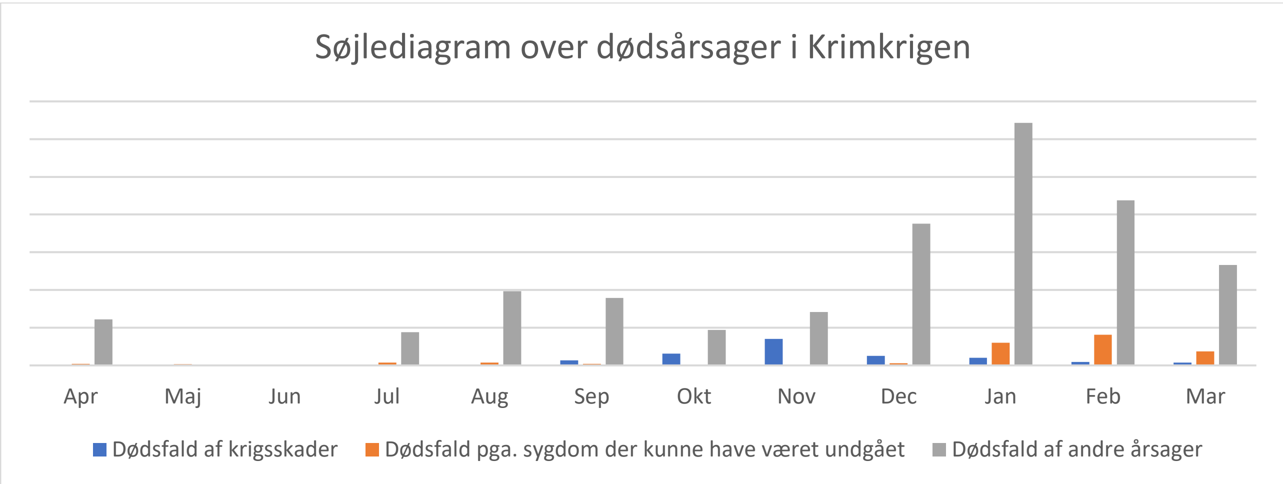
a. ”Cirklen” er inddelt i 12 sektorer, én for hver måned.
Antallet af dødsfald vokser kraftigt i juli og august og er højt flere måneder, men stiger yderligere hen over vintermånederne.

b. Man kan vurdere forholdet mellem arealet af de lyserøde og lyseblå områder. For at finde disse arealer kunne man måle radius for hvert område i en bestemt sektor, beregne arealet af de tilhørende cirkler, og dividere med 12, da vi ved, at diagrammet er inddelt i 12 sektorer.

c. Radius for hver sekter er målt i cm og angivet i følgende tabel. Desuden er arealet af hvert område beregnet.

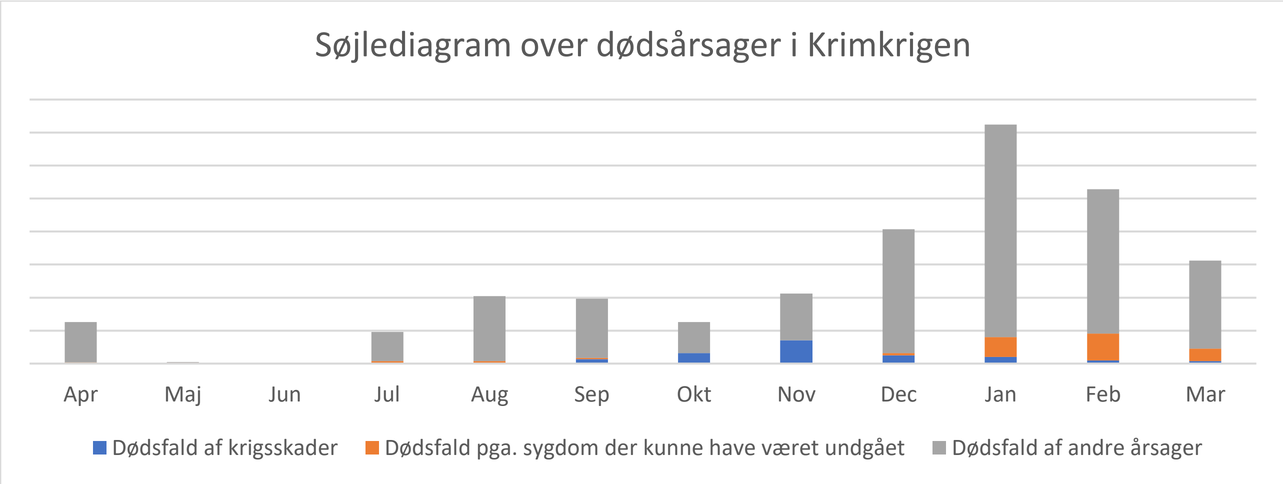
Måned	Apr	Maj	Jun	Jul	Aug	Sep	Okt	Nov	Dec	Jan	Feb	Mar
r lyserød	0	0	0	0	0	0,65	1	1,5	0,9	0,8	0,55	0,5
r sort	0,35	0,3	0,1	0,5	0,5	0,75	1	1,5	1	1,6	1,7	1,2
r lyseblå	2	0,4	0,1	1,75	2,55	2,5	2	2,6	3,6	4,8	4,1	3,15
A lyserød	0,00	0,00	0,00	0,00	0,00	1,33	3,14	7,07	2,54	2,01	0,95	0,79
A sort	0,38	0,28	0,03	0,79	0,79	0,44	0,00	0,00	0,60	6,03	8,13	3,74
A lyseblå	12,18	0,22	0,00	8,84	19,64	17,87	9,42	14,17	37,57	64,34	43,73	26,65

Herunder er arealerne afbilledet som et søjlediagram.

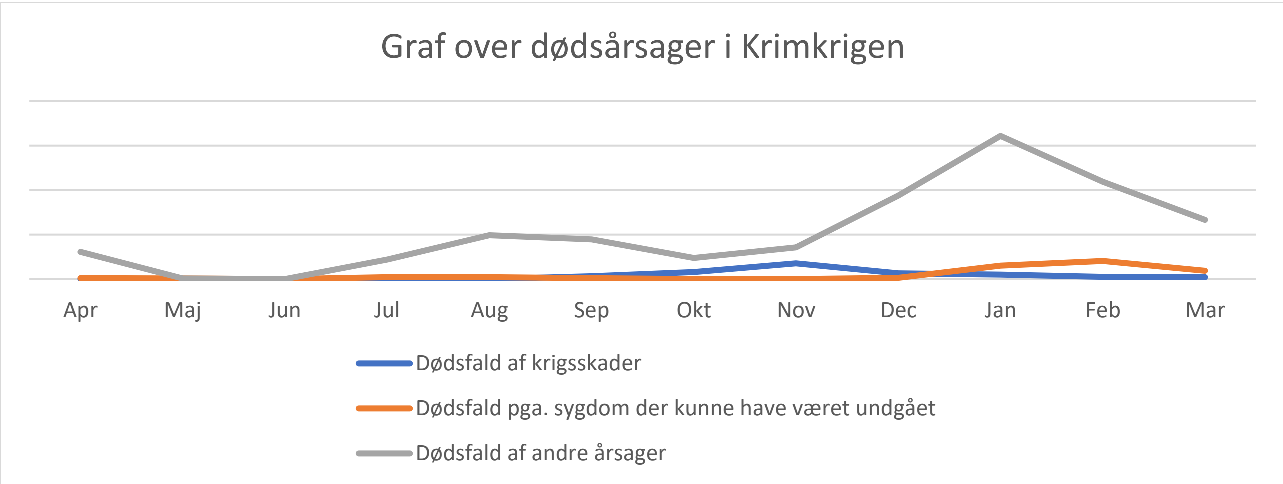


ISBN 9788770668279

Og som et stablet søjlediagram.



d.



L&R Uddannelse EGMONT

Øvelse 2.4

a. Et eksempel er datasættet $\{2,3,5,7\}$. Middeltallet bliver 4,25. Det vil det faktisk altid være, idet et generelt datasæt, der opfylder kravene, er på formen $\{2, y, z, 7\}$ hvor y og z er to tal mellem 2 og 7 der opfylder $\frac{y+z}{2}=4$ dvs. $y+z=8$ og dermed er middeltallet $\frac{2+y+z+7}{4}=\frac{2+8+7}{4}=\frac{17}{4}=4,25$.

b. Et eksempel er datasættet $\{2,3,4,7\}$. Medianen bliver 3,5. Det vil den faktisk altid være, idet et generelt datasæt, der opfylder kravene, er på formen $\{2, y, z, 7\}$ hvor y og z er to tal mellem 2 og 7 der opfylder $\frac{2+y+z+7}{4}=4$ dvs. $y+z=7$ og dermed er medianen $\frac{y+z}{2}=\frac{7}{2}=3,5$.

Øvelse 2.5

a. Middeltallet er 3.

b. Afstandene til middeltallet er angivet i følgende tabel:

Tal	1	1	2	5	6
Afstand til middelværdien	2	2	1	2	3

c. Tallene der er mindre end middeltallet er 1, 1 og 2. De har en samlet afstand på 5 til middeltallet. Tallene der er større end middeltallet er 5 og 6. De har også en samlet afstand på 5 til middeltallet. Dermed har vi kontrolleret, at middeltallet er balancepunkt for afstandene.

Øvelse 2.6

Middeltallet ændres i samme retning som observationen flyttes i, mens medianen er upåvirket af observationens værdi.

Øvelse 2.7

a. Middeltallet og medianen er begge 5.

b. Middeltallet er $125,\overline{3}$ og medianen er 5.

ISBN 9788770668279

Øvelse 2.8

a. $\{1,4,7,8,9,13,14,20,21,25\}$.

b. $\{1,4,7,8,9,13,14,20,21,2500\}$.

Øvelse 2.11

a. Medianen bliver også fordoblet (eller tredoblet).

b. Hvis alle observationerne i et datasæt ganges med en konstant k, så vil medianen for datasættet også blive ganget med den samme konstant k. Hvis M_0 er medianen for det oprindelige datasæt, og M er medianen for det opskalerede datasæt, så gælder $M = k \cdot M_0$.

c. Fattigdomsgrænsen er halvdelen af medianindkomsten, så når medianen bliver dobbelt så stor, bliver fattigdomsgrænsen også dobbelt så stor.

d. Der vil være lige så mange med en indkomst under fattigdomsgrænsen, så antallet af fattige ændres ikke, når alle indkomster fordobles.

Øvelse 2.12

a+b) Medianen og middeltallet er 50 for begge datasæt. Disse to statistiske deskriptorer, siger altså ikke noget om, hvor langt observationerne ligger fra hinanden.

Øvelse 2.15

a. $Q_1 = 1$, $Q_2 = 2$, $Q_3 = 5{,}5$.

b. Kvartilbredden er 4,5.

ISBN 9788770668279

Øvelse 2.16

a. Et eksempel er $\{1, 2, 4, 6, 9, 12\}$.

b. Middeltallet er $\frac{17}{3} = 5,\bar{6}$. Det vil det faktisk altid være, idet et generelt datasæt der opfylder kravene, er på formen $\{1, 2, y, z, 9, 12\}$ hvor y og z er to tal mellem 2 og 9 der opfylder $\frac{y+z}{2} = 5$ dvs. $y + z = 10$ og dermed er middeltallet $\frac{1+2+y+z+9+12}{6} = \frac{3+10+21}{6} = \frac{34}{6} = \frac{17}{3}$.

Øvelse 2.17

Størrelse af datasæt	5	6	7	8	9	10	11	12	13	14
Minimalt antal elementer i $[Q_1, Q_3]$	3	4	5	4	5	6	7	6	7	8

Det minimale antal elementer i intervallet stiger med 1 hver gang der tilføjes et element til datasættet, dog går det ned med 1, hvis den nye størrelse af datasættet er delelig med 4. Vi kan opskrive en formel for dette. Størrelsen af datasættet kan altid skrives som $4n + r$ hvor n er et helt tal og r er et helt tal fra 0 til 3. Da er det minimale antal elementer i $[Q_1, Q_3]$ givet ved $2n + r$.

Øvelse 2.19

a. $Q_1 = 4$, $Q_2 = 8,5$, $Q_3 = 10$. Kvartilbredden er 6.

b. Den mindste karakter er 2 og den største er 12. Medianen er gennemsnittet af de to midterste karakterer, som derfor må være 7 og 10. Første kvartil er gennemsnittet af 2. og 3. karakter, så det må være to 4-taller. Ligeledes er tredje kvartil gennemsnittet af 6. og 7. karakter, så det må være to 10-taller. Vi får derfor datasættet $\{2, 4, 4, 7, 10, 10, 10, 12\}$.

Øvelse 2.20

Vi tænker os, at karaktererne sættes i rækkefølge efter størrelse. Medianen er den 9. karakter i rækken. Nedre kvartil er gennemsnittet af den 4. og den 5. karakter, som begge må være 4. Øvre kvartil er gennemsnittet af den 13. og 14. karakter, som begge må være 10. Karaktersættet ser nu foreløbigt således ud: $\{2,*,*,4,4,*,*,*,7,*,*,*,10,10,*,*,12\}$, hvor *'erne står for ukendte karakterer. For at få det størst mulige gennemsnit, skal disse ukendte karakterer sættes så højt som muligt, og for at få det mindst mulige gennemsnit, skal de ukendte karakterer sættes så lavt som muligt. Dette giver:

a. Karaktersættet $\{2,4,4,4,4,7,7,7,7,10,10,10,10,10,12,12,12\}$ har det størst mulige gennemsnit på 7,76.

b. Karaktersættet $\{2,2,2,4,4,4,4,4,7,7,7,7,10,10,10,10,12\}$ har det mindst mulige gennemsnit på 6,24.

Øvelse 2.21

a. $\{1,2,3,4,5,6,7,8,9\}$.

b. De bliver begge 5. For et symmetrisk datasæt gælder, at medianen og middelværdien altid er lig den værdi, som datasættet er symmetrisk omkring.

c. $\{1,3,4,6,8,8\}$.

Øvelse 2.22

Nedre kvartil er 58 og kvartilbredden er 11. Så observationer mindre end $58-11\cdot 1,5=41,5$ er outliers. Derfor er 31 en outlier.

Øvelse 2.23

Værdier under -7 eller over 25 er outliers.

Øvelse 2.28

Ofrets alder	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-	I alt
Antal (hyppighed)	419	8176	10553	7175	5452	4274	2475	1995	1399	41918
Procent (frekvens)	1,00	19,50	25,18	17,12	13,01	10,20	5,90	4,76	3,34	100

ISBN 9788770668279

Øvelse 2.29

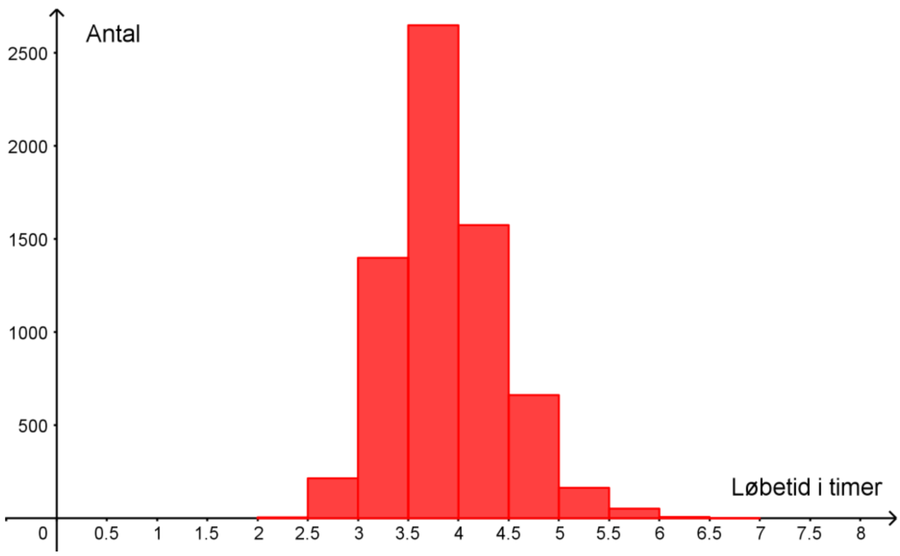
b. Medianen må befinde sig i intervallet 30-39 år, da det er her den midterste observation befinder sig.

Øvelse 2.30

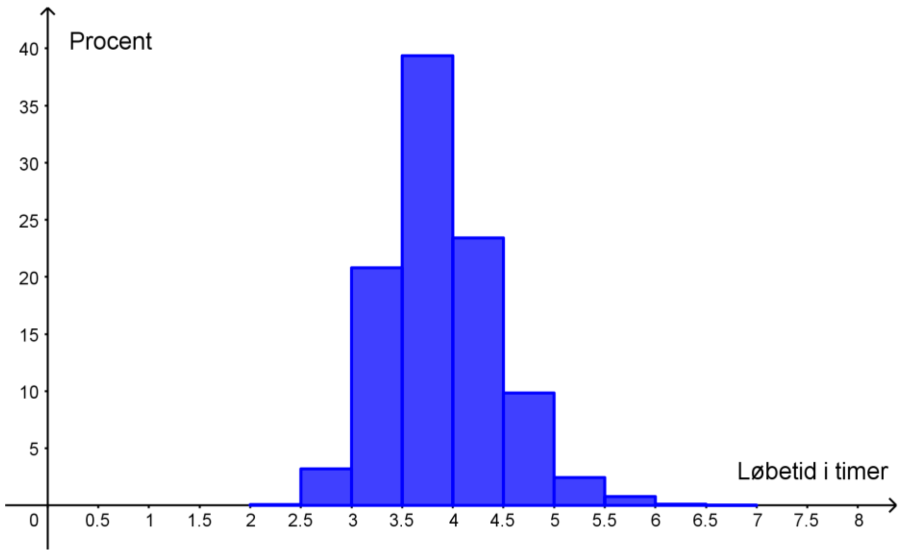
a.

Løbetid	2:00-2:30	2:30-3:00	3:00-3:30	3:30-4:00	4:00-4:30	4:30-5:00	5:00-5:30	5:30-6:00	6:00-6:30	6:30-7:00	I alt
Antal	5	215	1399	2648	1575	662	164	52	7	1	6728
Procent	0,074	3,196	20,794	39,358	23,410	9,839	2,438	0,773	0,104	0,015	100

b. Medianen må befinde sig i intervallet 3:30-4:00, da det er her den midterste observation befinder sig.



ISBN 9788770668279



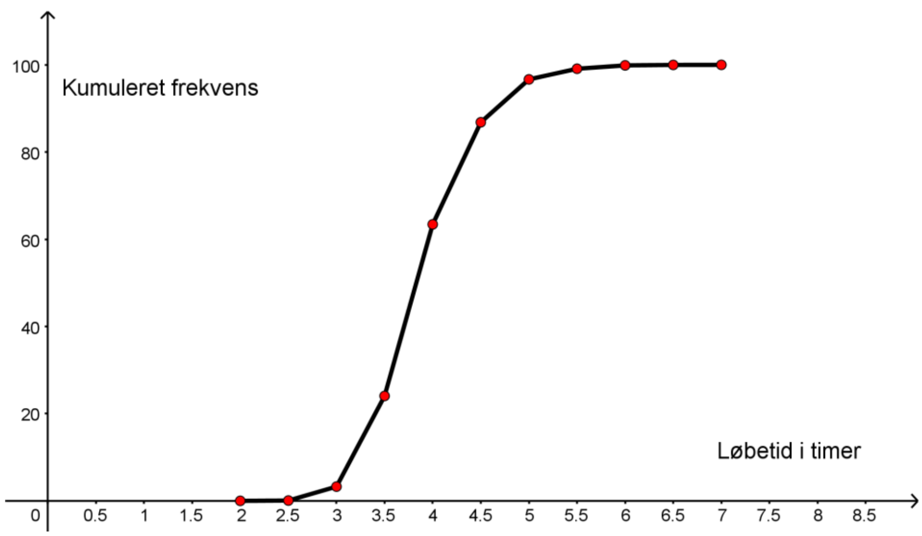
Øvelse 2.32

3,88 timer.

Øvelse 2.35

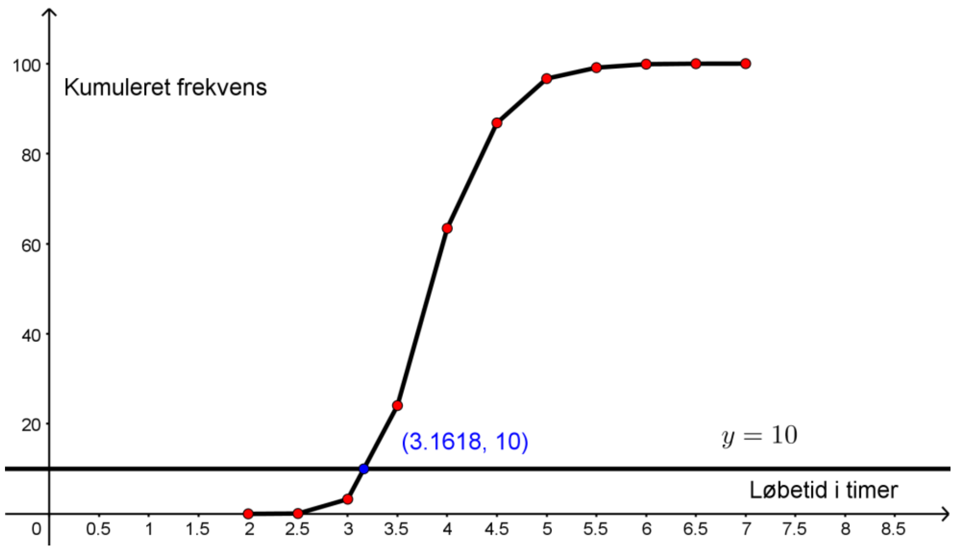
a.

Løbetid	2:00-2:30	2:30-3:00	3:00-3:30	3:30-4:00	4:00-4:30	4:30-5:00	5:00-5:30	5:30-6:00	6:00-6:30	6:30-7:00
Antal	5	215	1399	2648	1575	662	164	52	7	1
Procent	0,074	3,196	20,794	39,358	23,410	9,839	2,438	0,773	0,104	0,015
Kumuleret Frekvens	0,074	3,270	24,064	63,422	86,831	96,671	99,108	99,881	99,985	100,00



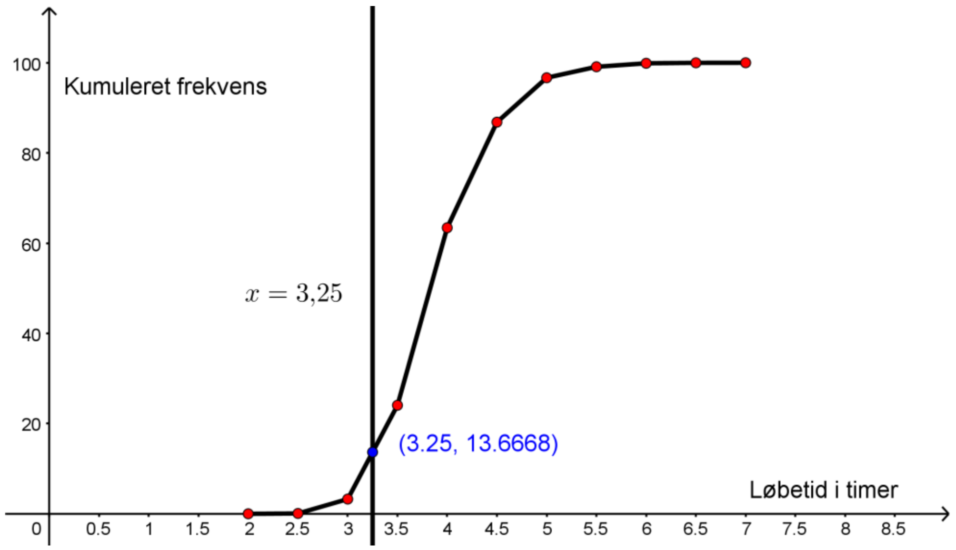
ISBN 9788770668279

b. Vi bestemmer grafisk skæringspunktet med linjen $y = 10$.



De 10% hurtigste løbere bruger under 3,16 timer.

c. Vi finder skæringspunktet med linjen $x = 3{,}25$ (da 15 minutter er en kvart time).



13,7% af løberne bruger under 3,25 timer (3 timer og 15 minutter).

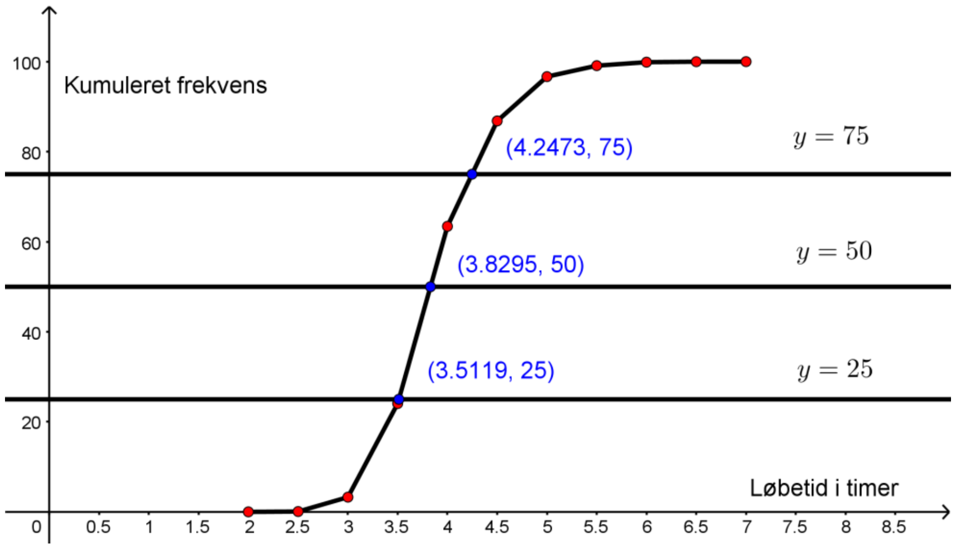
d. Sumkurven er stejlest i intervallet 3,5-4,0.

e. Det interval hvor sumkurven er stejlest, er der den største procentdel af løbetiderne befinder sig.

ISBN 9788770668279

Øvelse 2.36

a.



Kvartilsættet er $Q_1 = 3{,}51$, $Q_2 = 3{,}83$, $Q_3 = 4{,}25$ (alle tre målt i timer).

b. Middeltallet er 3,88 timer. Det er meget tæt på medianen, så fordelingen af løbetider er ganske symmetrisk, dog med en lille andel af store løbetider, som trækker middelværdien ganske lidt op. De 25% hurtigste løber på under tre en halv time. De 25% langsomste løber på over fire timer og et kvarter. Halvdelen af løberne kommer altså i mål indenfor et tidsinterval på 45 minutter (kvartilbredden). Pga. den lille kvartilbredde er der 60 outliere, som er de løbere der bruger over fem en halv time. Herudover kan enkelte løbere i intervallerne 5:00-5:30 og 2:00-2:30 være outliere.

Øvelse 2.37

Status\Skæbne	Overlevede	Omkom	I alt
Første	199	130	329
Anden	119	166	285
Tredje	174	536	710
I alt	492	832	1324

Øvelse 2.38

Status\Skæbne	Overlevede	Omkom	I alt
Første	60,49%	39,51%	100%
Anden	41,75%	58,25%	100%
Tredje	24,51%	75,49%	100%
I alt	37,16%	62,84%	100%

ISBN 9788770668279

Øvelse 2.39

a. 62,84% af alle ombord omkom.

b. 75,49%.

c. 58,25%.

d. 39,51%.