

EKSAMEN IRBIO10620 – VÅR 2021 – LØSNINGSFORSLAG

Løsningsforslaget tilsvarer karakteren A.

Oppgave 1

Oppgaven gir 5 poeng.

Vi har en populasjon på 5 millioner personer.

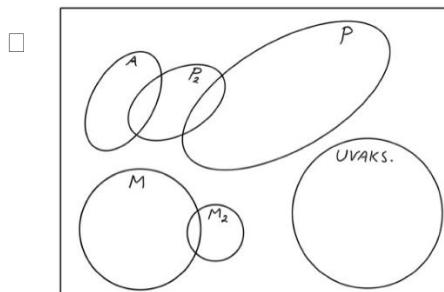
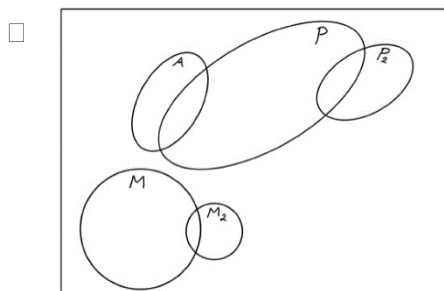
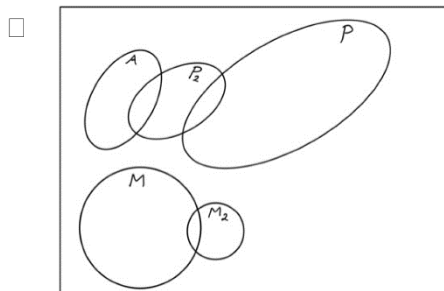
180000 personer har fått én dose AstraZeneca (**A**).

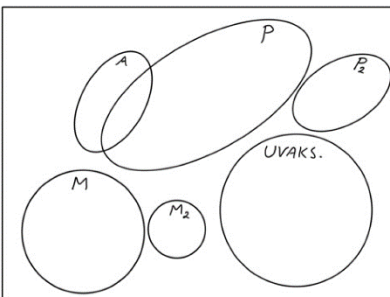
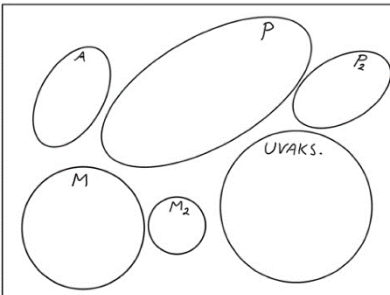
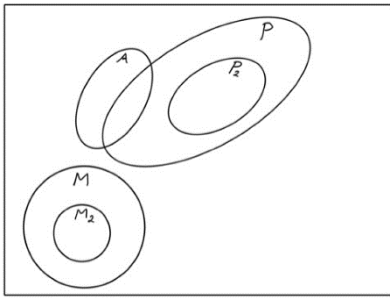
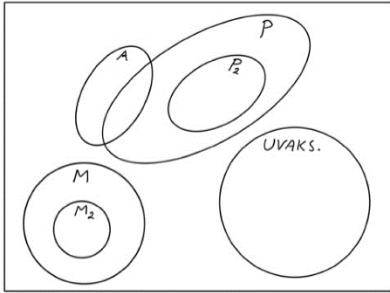
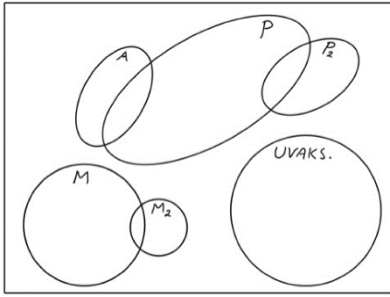
300000 personer har fått kun én dose Moderna (**M**), 35000 har fått to doser Moderna (**M₂**).

2000000 personer har kun fått én dose Pfizer (**P**) og ingenting annet, 70000 har fått én dose Pfizer i tillegg til en **A**, 220000 har fått to doser Pfizer (**P₂**).

Hvilket Venndiagram illustrerer denne beskrivelsen?

Velg ett alternativ:





Oppgave 2

Oppgaven gir 5 poeng.

Vi har en populasjon på 5 millioner personer.

180000 personer har fått én dose AstraZeneca (**A**).

300000 personer har fått kun én dose Moderna (**M**), 35000 har fått to doser Moderna (**M**₂).

2000000 personer har kun fått én dose Pfizer (**P**) og ingenting annet, 70000 har fått én dose Pfizer i tillegg til en **A**, 220000 har fått to doser Pfizer (**P**₂).

Hvordan vil du skrive den mengden av personer som kun har fått én dose og at den var av type AstraZeneca

Velg ett alternativ:

- $A \cup \bar{P}$
- Ingen av alternativene er riktig
- $A \cup P$
- $\bar{A} \cup P$
- $\bar{A} \cup \bar{P}$
- $A \cap \bar{P}$
- $\bar{A} \cap \bar{P}$
- $A \cap P$
- $\bar{A} \cap P$

Hint til løsning: Har fått A og samtidig ikke fått P.

Oppgave 3

Oppgaven gir 5 poeng.

Vi har 169 ulike tall, variansen av disse er 25. Hva er standardavviket til tallene?

Velg ett alternativ:

- 65
- 5
- 13
- 25
- 4225
- Ingen av alternativene er riktig
- 28561
- 625

Hint til løsning: Kvadratrot av variansen.

Oppgave 4

Oppgaven gir 5 poeng.

En vaksine gir 95% beskyttelse mot en gitt sykdom. 9 personer får vaksinen og utsettes senere for smitte. Hva er sannsynligheten for at ingen av personene blir sjuke?

Velg ett alternativ:

- Ingen av alternativene er riktig.
- 0,370
- 0,630
- 0,663
- 0,337
- 0,086

Hint til løsning: Binomisk fordeling, $p=0.05$, $n=9$, $x=0$

Oppgave 5

Oppgaven gir 5 poeng.

På et laboratorium analyseres det daglig 9 prøver. Vi vet at 1 av analysene er gale, men det vet ikke noen på laboratoriet. Et utvalg på 4 av de 9 prøvene analyseres på nytt. Hvor sannsynlig er det at utvalget ikke inneholder prøven som var feil ved første analyse?

Velg ett alternativ:

- Ingen av alternativene er riktige
- 0,333
- 0,667
- 0,556
- 0,778
- 0,444

Hint til løsning: Hypergeometrisk fordeling, én enhet med bestemt egenskap.

Oppgave 6

Oppgaven gir 5 poeng.

Etter mange års erfaring forventer Kaosklinikken at det kommer 12 pasienter hver time. (Døra til Kaosklinikken er så smal at det kommer kun én pasient av gangen.) Hva er sannsynligheten for at det ikke kommer noen pasienter på 1 time?

Velg ett alternativ:

- 0,06
- 0,006
- Ingen av alternativene er riktige.
- 0,000006
- 0,0006
- 0,6

Hint til løsning: Poissonfordeling.

Oppgave 7

Oppgaven gir 5 poeng.

Kaosklinikken kjøper vann på flasker. I følge produsenten er innholdet i hver flaske 1000 mL med standardavvik 1 mL, vi antar at volumet er normalfordelt. Hva er sannsynligheten for at mengden vann i en flaske er mer enn 998 mL?

Velg ett alternativ:

- 0,9309
- 0,9971
- 0,9653
- 0,9487
- 0,9772
- Ingen av alternativene er riktige.

Studentene fikk oppgitt følgende formel på eksamen:

Regel 5.14 Normalfordelingens fordelingsfunksjon

La X være Normal(μ, σ). Da gjelder:

$$P(X \leq x) = F(x) = G\left(\frac{x - \mu}{\sigma}\right)$$

$$P(X > x) = 1 - G\left(\frac{x - \mu}{\sigma}\right)$$

$$P(a \leq X < b) = G\left(\frac{b - \mu}{\sigma}\right) - G\left(\frac{a - \mu}{\sigma}\right)$$

Hint til løsning: Normalfordeling.

Oppgave 8

Oppgaven gir 5 poeng.

I trafikken er det en øvre grensen for hvor mye alkohol en bilfører kan ha i blodet. Er alkoholnivået over grensen vil bilføreren få bot eller fengsel. For å avgjøre om bilføreren har gjort noe galt eller ikke kan vi bruke en hypotesetest. Hvilke av følgende hypoteser skal vi velge?

Alkoholnivået bestemmes ut fra en blodprøve. Analysen gjøres med en vel etablert metode. Når analyseresultatet foreligger må vi velge hypotesetest for å avgjøre om nullhypotesen skal forkastes. Vi skal da velge følgende test:

Hint til løsning: Hypotesetest, konservativ H0 er at fører er lovlydig. Kjent standardavvik betyr Z-test.

Oppgave 9

Oppgaven gir 5 poeng.

Gjennomsnittlig IQ i befolkningen er 100 og tilnærmet normalfordelt. Vi liker å anta at bioingeniører har høyere IQ enn befolkningen ellers. En IQ-test på 9 bioingeniørstudenter gav resultatene: 103, 105, 108, 102, 107, 99, 101, 106 og 103. (Gjennomsnittet av dette er 103,8, standardavviket 2,95.)

Kan vi utfra disse tallene, med signifikansnivå 0.05 (alfa), si at bioingeniører har høyere IQ enn resten av befolkningen?

Velg ett alternativ:

- Ingen av testene kan brukes på denne type data.
- Må bruke en paret T-test, den gir oss at bioingeniører har høyere IQ.
- Må bruke en T-test, den gir oss at bioingeniører har høyere IQ.
- Må bruke en paret T-test, den gir oss at bioingeniører ikke har høyere IQ.
- Må bruke en T-test, den gir oss at bioingeniører ikke har høyere IQ.
- Må bruke en Z-test, den gir oss at bioingeniører har høyere IQ.
- Må bruke en Z-test, den gir oss at bioingeniører ikke har høyere IQ.

Hint til løsning: Hypotesetest, konservativ H_0 er at $\text{bioing} \leq 100$. Ukjent standardavvik betyr T-test.

Oppgave 10

Oppgaven gir 5 poeng.

Vi har følgende datasett bestående av ni datapunkter:

x	y
0	0
1	-1
1	1
2	-2
2	0
2	2
3	-1
3	1
4	0

Finn korrelasjonen.

[Korrelasjonen kan bestemmes enten ved å tegne figur eller regnes ut slik:

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \cdot \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Velg ett alternativ:

- 0,5
- 0,5
- 1
- Ingen av alternativene er riktige.
- 1
- 0

Hint til løsning: Ikke mulig å trekke noen meningsfull rett linje gjennom punktansamlingen.

Oppgave 11

Oppgaven gir 3 poeng.

Hva vil det si at en analysemetode har god riktighet, og hvilken parameter brukes for å angi analysemetodens uriktighet?

Les svaralternativene NØYE. I svaret du velger skal begge utsagn være rette.

Velg ett alternativ:

- At en analysemetode har god riktighet, betyr at analysemetoden ikke har tilfeldige feil. For å angi en analysemetodes uriktighet brukes standardavvik og variasjonskoeffisient.
- At en analysemetode har god riktighet, betyr at analysemetoden ikke har tilfeldige feil. For å angi en analysemetodes uriktighet brukes bias.
- At en analysemetode har god riktighet, vil si at dersom en måler analytten i samme prøvemateriale (referanseløsning) flere ganger, vil gjennomsnittet av disse resultatene være tilnærmet lik sann verdi av analytten. For å angi en analysemetodes uriktighet brukes bias.
- At en analysemetode har god riktighet, vil si at dersom en måler analytten i samme prøvemateriale (referanseløsning) flere ganger, vil gjennomsnittet av disse resultatene være tilnærmet lik sann verdi av analytten. For å angi en analysemetodes uriktighet brukes standardavvik og variasjonskoeffisient.
- At en analysemetode har god riktighet, betyr at analysemetoden ikke har systematiske feil. For å angi en analysemetodes uriktighet brukes standardavvik og variasjonskoeffisient.

Oppgave 12

Oppgaven gir 1 poeng.

Avgjør om følgende påstand er sann eller usann:

Falsk alarm betyr at man får en alarm når det ikke foreligger feil, dvs. når analysemetoden er under kontroll.

Velg ett alternativ:

- Sant
- Usant

Oppgave 13

Oppgaven gir 1 poeng.

Avgjør om følgende påstand er sann eller usann:

Et styrkeidiagram viser 1) sammenhengen mellom størrelsen på en analysefeil og sannsynlighet for at kontrollregelen gir alarm, P_{ed} (eller P_a) og 2) sannsynligheten for falsk alarm, P_{fr} (eller P_{fa}), ved bruk av kontrollregelen.

Velg ett alternativ:

Sant

Usant

Oppgave 14

Oppgaven gir 1 poeng.

Avgjør om følgende påstand er sann eller usann:

Hvis Shewhart-diagrammet viser at alle kontrollresultatene trekkes i én retning og ser ut til å samle seg om en ny målverdi, har det oppstått en systematisk feil i analysemetoden.

Velg ett alternativ:

Sant

Usant

Oppgave 15

Oppgaven gir 1 poeng.

Avgjør om følgende påstand er sann eller usann:

De fleste kommersielle kontrollmaterialer har kortere holdbarhet enn egenproduserte kontrollmaterialer.

Velg ett alternativ:

Sant

Usant

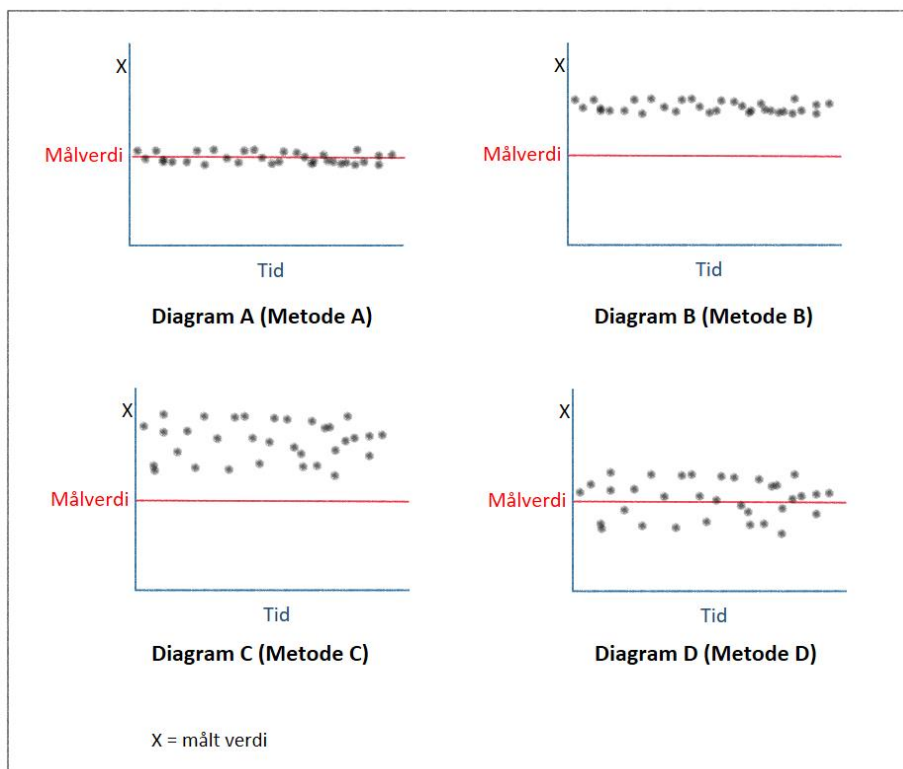
Oppgave 16

Oppgaven gir 3 poeng.

Diagrammene viser kontrollresultater for 4 metoder:

- Diagram A viser kontrollresultater for metode A.
- Diagram B viser kontrollresultater for metode B.
- Diagram C viser kontrollresultater for metode C.
- Diagram D viser kontrollresultater for metode D.

Hvilket av Shewhart-diagrammene viser en analysemetode med god riktighet, men dårlig presisjon?



Velg ett alternativ:

- Diagram A.
- Diagram B.
- Diagram C.
- Diagram D.
- Ingen av diagrammene.

Hint til løsning:

Dårlig presisjon: stor spredning mellom enkeltresultatene.

Dårlig riktighet: resultatene sprer seg rundt en annen middelvei enn målverdien.

Oppgave 17

Oppgaven gir 10 poeng.

Shewhart-diagrammet på bildet (Vedlegg til oppgave 17) viser kontrollresultater for måling av s-totalprotein i august 2020.

a)

Beskriv kort hva som skal til for å få alarm med den kombinerte kontrollregelen $1_{3s}/2_{2s}$.

(4 poeng.)

b)

Vurder kontrollresultatene ut fra den kombinerte kontrollregelen $1_{3s}/2_{2s}$, og bruk kontrollregelen 1_{2s} , som varsel:

1) Hvilke dager har laboratoriet fått varsel og hvorfor?

2) Hvilke dager har laboratoriet fått alarm?

For dager hvor laboratoriet har fått alarm: Skriv hvilken del av den kombinerte kontrollregelen som brytes, hvilken alarmgrense som brytes og hvorfor.

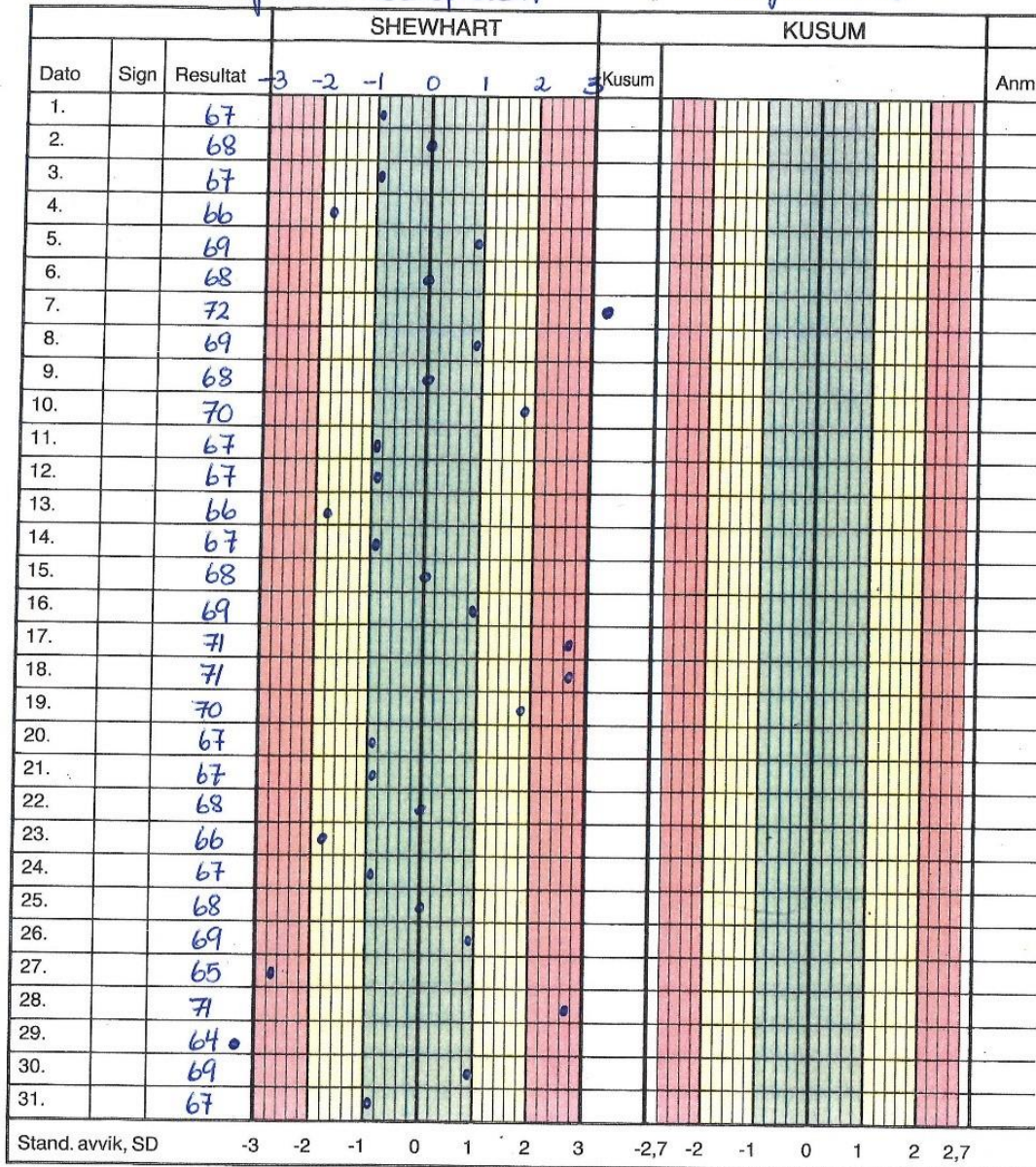
Svar kort. (6 poeng.)

Vedlegg til oppgave 17.

KONTROLLSKJEMA

Analyse: Måling av S-totalprotein

Periode: August 2020



Beregning: $\bar{x} = 68,0 \text{ g/L}$
 $s = 1,1 \text{ g/L}$

4113 - 2/88. MELSOM - 1652 TORP

Løsning:

a)

Den kombinerte kontrollregelen $1_{3s}/2_{2s}$ gir alarm hvis 1 kontrollresultat ligger utenfor målverdi $\pm 3s$, altså utenfor $-3s$ -grensen eller utenfor $+3s$ -grensen. Denne alarmregelen gir også alarm hvis 2 kontrollresultater etter hverandre ligger utenfor samme $2s$ -grense (dvs. begge kontrollresultater er utenfor målverdi- $2s$, eller begge kontrollresultater er utenfor målverdi+ $2s$).

b)

Laboratoriet har fått varsel fordi 1_{2s} -regelen brytes på følgende dager:

- på dag 17 og dag 28: fordi ett kontrollresultat er utenfor $+2s$ -grensen (målverdi + $2s$).
- på dag 27 fordi ett kontrollresultat er utenfor $-2s$ -grensen (målverdi- $2s$)

c)

Laboratoriet har fått alarm på dag 7, på dag 18 og på dag 29.

På dag 7 brytes kontrollregel 1_{3s} fordi ett kontrollresultat er utenfor $+3s$ -grensen (målverdi + $3s$).

På dag 18 brytes kontrollregel 2_{2s} fordi 2 påfølgende kontrollresultater er utenfor samme $2s$ -grense (målverdi+ $2s$).

På dag 29 blir det alarm fordi kontrollregelen 1_{3s} brytes pga. at ett kontrollresultat er utenfor $-3s$ -grensen (målverdi - $3s$).

Oppgave 18

Oppgaven gir 4 poeng.

Bruk vedlagte styrkeagrammer (Vedlegg til oppgave 18) til å finne sannsynligheten for å få alarm, P_{ed} , (eller P_a), med kontrollregel 1_{2s} hvis det oppstår en systematisk feil på 1,5 ganger standardavviket ($1,5s$) for analysemetoden.

Hva er sannsynlighet for falsk alarm, P_{fr} (eller P_{fa}), for denne kontrollregelen?

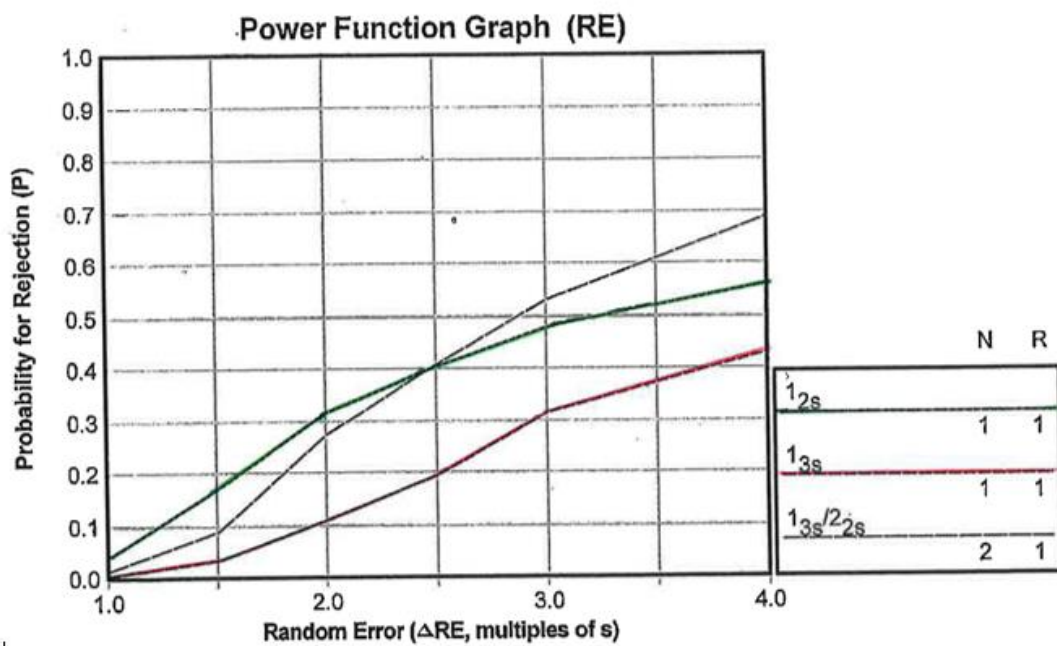
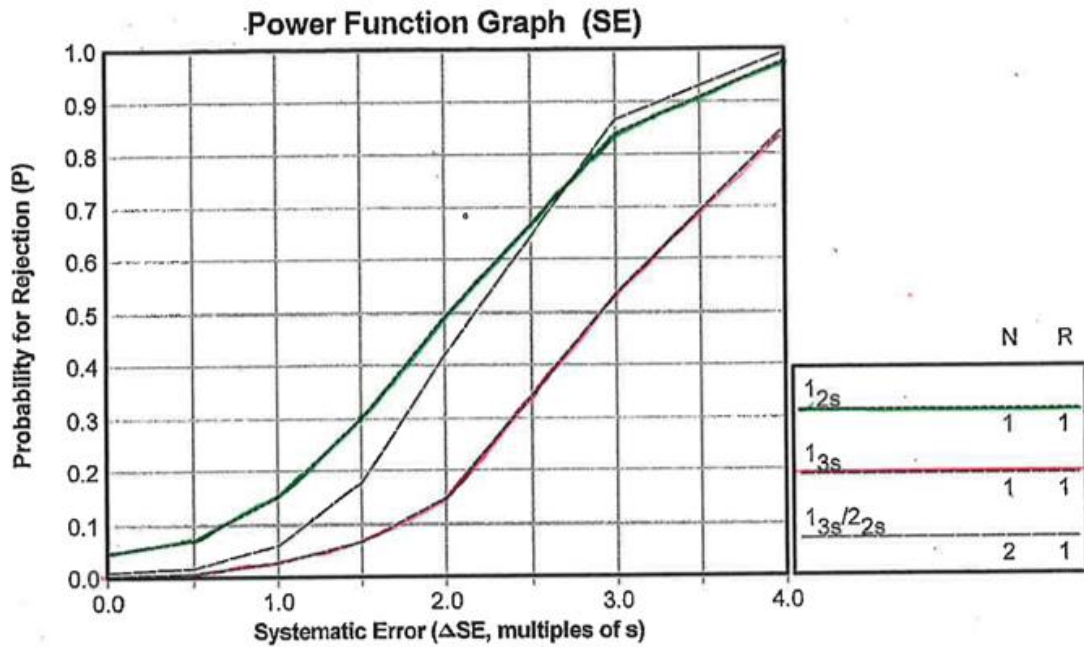
Les svaralternativene NØYE. I svaret du velger skal begge setninger være rette.

Velg ett alternativ:

- P_{ed} er ca. 0,07 (ca. 7 %), og P_{fr} er $< 0,01$ (< 1 %).
- P_{ed} er ca. 0,18 (ca. 18 %), og P_{fr} er ca. 0,02 (ca. 2 %).
- P_{ed} er ca. 0,30 (ca. 30 %), og P_{fr} er ca. 0,05 (ca. 5 %).
- P_{ed} er ca. 0,83 (ca. 83 %), og P_{fr} er ca. 0,05 (ca. 5 %).
- P_{ed} er ca. 0,30 (ca. 30 %), og P_{fr} er ca. 0,02 (ca. 2 %).

Vedlegg til oppgave 18.

Styrkediagrammer.



Hint til løsning:

$P_{ed}(SE)$ leses av ved $\Delta SE = 1,5$ for kontrollregel 1_{2s} , P_{fr} leses av ved $\Delta SE = 0$ for kontrollregel 1_{2s} .

Oppgave 19

Oppgaven gir 5 poeng.

En ny urintest for *Chlamydia trachomatis* ble utviklet, og følgende resultater ble oppnådd:

765 menn ble testet.

85 menn fikk positiv test, og av disse var 10 menn bekreftet friske.

680 menn fikk negativ test, og av disse var 18 menn bekreftet syke.

Diagnostisk sensitivitet er sannsynligheten for at en syk person får et positivt testresultat.

Hva blir diagnostisk sensitivitet (i prosent) for den nye testen?

Velg ett alternativ:

- Diagnostisk sensitivitet er 75 %.
- Diagnostisk sensitivitet er 89 %.
- Diagnostisk sensitivitet er 81 %.
- Diagnostisk sensitivitet er 97 %.
- Diagnostisk sensitivitet er 99 %.

Hint til løsning: Bruk formel for beregning av diagnostisk sensitivitet.

Oppgave 20

Oppgaven gir 4 poeng.

Se vedlagt tegning (Vedlegg til oppgave 20).

Hvilken av disse betegnelsene:

- Sanne negative
- Sanne positive
- Falske negative
- Falske positive

hører hjemme i hvilket felt på tegningen?

Velg ett alternativ:

- Felt 1: Falske negative
- Felt 2: Sanne negative
- Felt 3: Sanne positive
- Felt 4: Falske positive

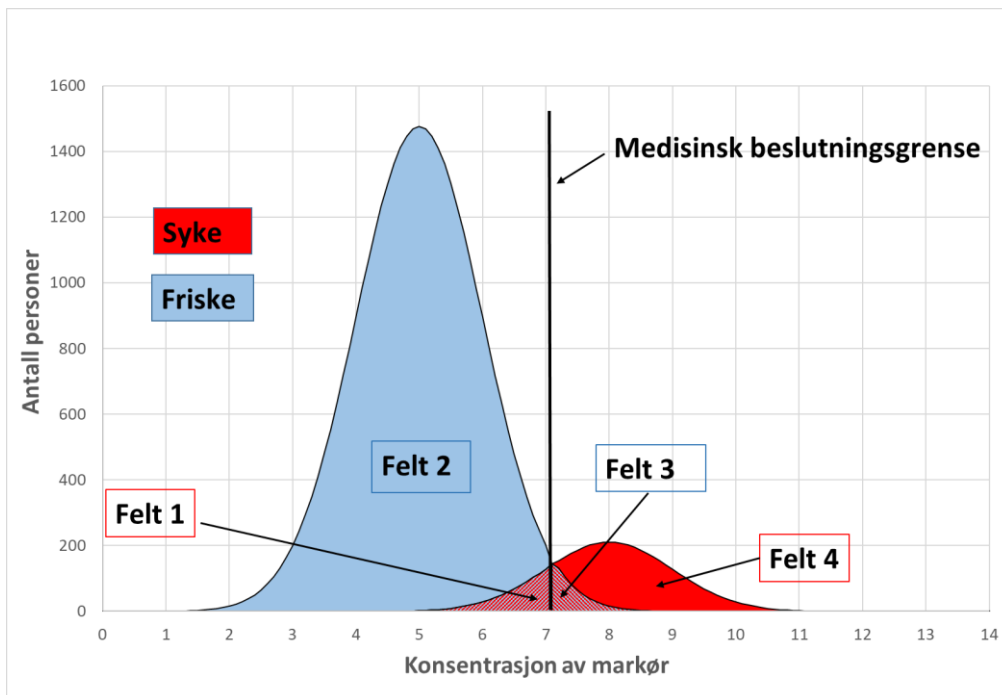
- Felt 1: Sanne negative
- Felt 2: Falske positive
- Felt 3: Falske negative
- Felt 4: Sanne positive

- Felt 1: Sanne positive
- Felt 2: Falske positive
- Felt 3: Falske negative
- Felt 4: Sanne negative

- Felt 1: Falske negative
- Felt 2: Sanne negative
- Felt 3: Falske positive
- Felt 4: Sanne positive

- Felt 1: Falske positive
- Felt 2: Sanne negative
- Felt 3: Falske negative
- Felt 4: Sanne positive

Vedlegg til oppgave 20.



Hint til løsning:

Felter på venstresiden av beslutningsgrensen (lavere verdier av analytt) gir negativt testresultat.

Felter på høyresiden av beslutningsgrensen (høyere verdier av analytt) gir positivt testresultat.

Felt 1: Syke som har fått negativt svar: Falske negative.

Felt 2: Friske som har fått negativt svar: Sanne negative.

Felt 3: Friske som har fått positivt svar: Falske positive.

Felt 4: Syke som har fått positivt svar: Sanne positive.

Oppgave 21

Oppgaven gir 5 poeng.

Under en validering av en analysemetode ble følgende verdier bestemt for analytt A:

Gjennomsnittskonsentrasjon av analytt A = 80,5 mg/L

Innen-serien variasjonen, $s_w = 1,5$ mg/L

Dag-til-dag variasjonen, $s_t = 4,0$ mg/L.

Beregn mellom-serie variasjonskoeffisienten, CV_b (kan også forkortes VK_b).

Velg ett alternativ:

- $CV_b = 2,9 \%$
- $CV_b = 5,3 \%$
- $CV_b = 2,0 \%$
- $CV_b = 17,1 \%$
- $CV_b = 4,6 \%$

Hint til løsning: Bruk formelen $s_t^2 = s_w^2 + s_b^2$, deretter formelen: $CV\% = 100 \cdot \frac{s}{\bar{x}}$.

Oppgave 22

Oppgaven gir 12 poeng.

I vedlegget (Vedlegg til oppgave 22) vises spredningsdiagram og regresjonsstatistikk fra en metodesammenligning hvor en gammel, anerkjent analysemetode (Metode A) sammenlignes med en ny analysemetode (Metode B) for analytt Y (g/L).

a)

Hva blir regresjonslinjen og korrelasjonskoeffisienten R? (2 poeng.)

Løsning:

Regresjonslinjen: $y = bx + a$, der $b = \text{stigningstallet}$ og
 $a = \text{skjæringspunktet (med y-aksen) for regresjonslinjen,}$

blir: $y = 1,03x - 1,54$

Korrelasjonskoeffisienten, $R = 0,986 \approx 0,99$

b)

Hva er 95 % konfidensintervall for 1) stigningstallet og 2) skjæringspunktet? (2 poeng.)

Løsning:

95 % konfidensintervall for stigningstallet b er: $[0,98, 1,08]$ (ingen benevning).

95 % konfidensintervall for skjæringspunktet a er: $[-3,93, 0,85]$ g/L.

c)

Er det et signifikant konsentrasjonsavhengig systematisk avvik og/eller et signifikant konsentrasjonsuavhengig systematisk avvik mellom metodene på 5 % signifikansnivå?

Begrunn svaret.

(Hvis du ikke klarer oppgave a) og b), så beskriv i detalj hvordan du ville løst oppgave c) hvis du hadde klart oppgave a) og b.)

(4 poeng.)

Løsning:

Når metodene er like, er $y = x$, da er stigningstallet $b = 1$ og skjæringspunktet $a = 0$. Den ideelle verdien for stigningstallet b er altså 1, og den ideelle verdien for skjæringspunktet a er 0.

Siden den ideelle verdien 1 er med i 95 % konfidensintervall ([0,98, 1,08]) for stigningstallet b , kan det ikke påstås at stigningstallet b er signifikant forskjellig fra 1 (på 5 % signifikansnivå).
Det er derfor ikke noe signifikant konsentrasjonsavhengig systematisk avvik mellom de to metodene på 5 % signifikansnivå.

Den ideelle verdien for skjæringspunktet a er 0.

Denne verdien ligger i 95 % konfidensintervall ([-3,93, 0,85] g/L) for skjæringspunktet a , og vi kan derfor ikke påstå at skjæringspunktet a er signifikant forskjellig fra 0 (på 5 % signifikansnivå).

Det er derfor ikke noe signifikant konsentrasjonsuavhengig systematisk avvik mellom metodene på 5 % signifikansnivå.

d)

Beregn bias i % (bias%) mellom analysemetode A og analysemetode B ved beslutningsgrensen 30 g/L ved å bruke regresjonslinjen. Forklar i detalj hva du gjør og vis alle beregninger.

(Hvis du ikke klarer oppgave a), så beskriv i detalj hvordan du ville løst oppgave d) hvis du hadde klart oppgave a)).

(4 poeng.)

Løsning:

Regresjonsligningen er: $y = 1,03x - 1,54$, der stigningstallet $b = 1,03$, og skjæringspunktet $a = -1,54$ g/L.

Bias skal beregnes ved beslutningsgrensen 30 g/L.

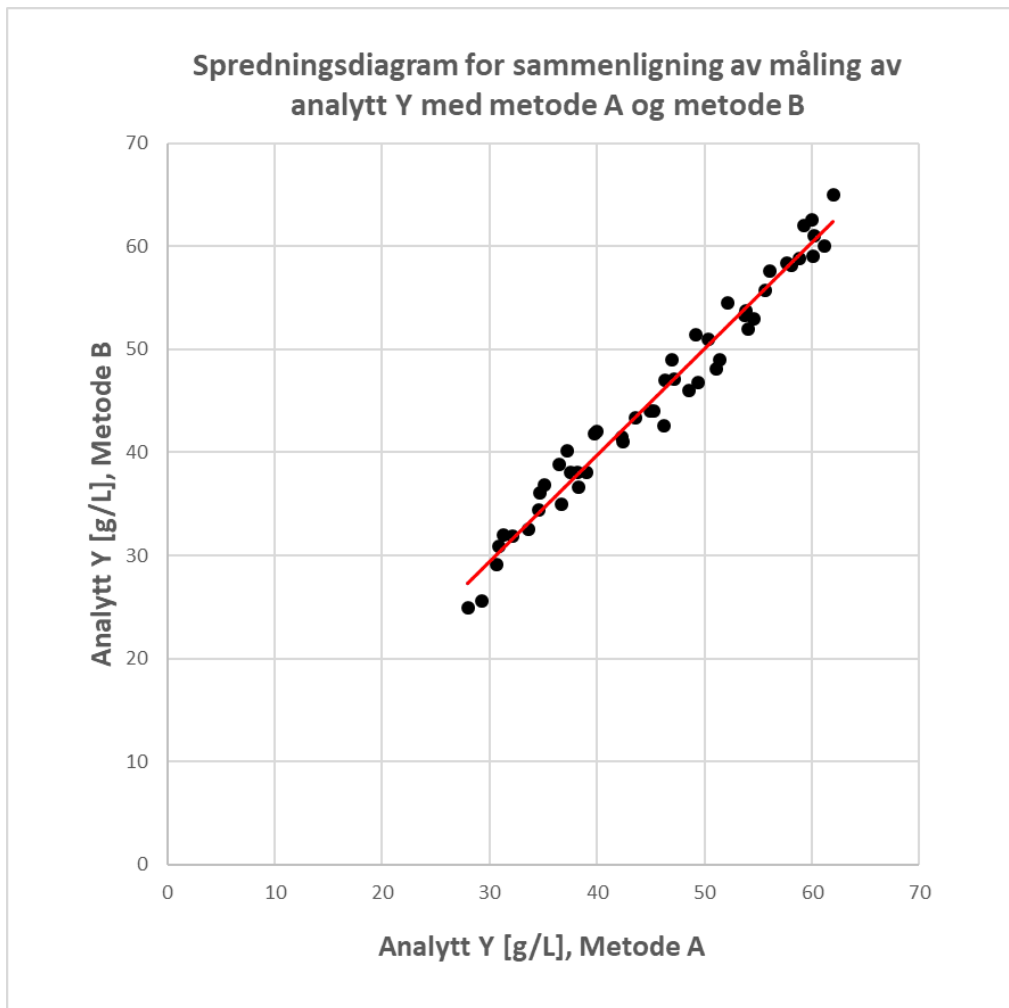
Dette gjøres ved å sette 30 g/L inn for x i regresjonsligningen og finne tilsvarende y -verdi.

(x = Metode A og y = metode B (jfr. spredningsdiagrammet).)

$$y = (1,03 \cdot 30 - 1,54) \text{ g/L} = 29,36 \text{ g/L}$$

$$\text{Bias \%} = 100 \% \cdot (y-x)/x = 100 \% \cdot (29,36-30)/30 = -2,13 \%$$

Vedlegg til oppgave 22.



Figur 1. Spredningsdiagram for sammenligning av måling av analytt Y (g/L) med metode A og metode B.

Tabell 1. Regresjonsstatistikk fra lineær regresjon med minste kvadraters metode ved sammenligning av Metode A og Metode B for analytt Y (g/L).

<i>Regresjonsstatistikk</i>	
Multipel R	0,985706585
R-kvadrat	0,971617472
Justert R-kvadrat	0,971026169
Standardfeil	1,782555235
Observasjoner	50

	<i>Koeffisienter</i>	<i>Standardfeil</i>	<i>Nederste 95%</i>	<i>Øverste 95%</i>
Skjæringspunkt	-1,540098611	1,189490393	-3,93172934	0,851532118
Stigningstall	1,030531833	0,025422525	0,979416422	1,081647245