i Forside mai 2021

Institutt for matematiske fag

Eksamensoppgave i TMA4245 Statistikk

Eksamensdato: Fredag 14. mai, 2021

Eksamenstid (fra-til): 09.00 – 13.00

Hjelpemiddelkode/Tillatte hjelpemidler: A / Alle hjelpemidler tillatt

Faglig kontakt under eksamen: Geir-Arne Fuglstad og Jarle Tufto
Tlf.: 45270806, 99705519

Teknisk hjelp under eksamen: NTNU Orakel
Tlf: 73 59 16 00

Får du tekniske problemer underveis i eksamen, må du ta kontakt for teknisk hjelp snarest mulig, og senest innen eksamenstida løper ut. Kommer du ikke gjennom umiddelbart, hold linja til du får svar.

ANNEN INFORMASJON:

Gjør dine egne antagelser og presiser i besvarelsen hvilke forutsetninger du har lagt til grunn i tolkning/avgrensing av oppgaven. Faglig kontaktperson skal kun kontaktes dersom det er direkte feil eller mangler i oppgavesettet.

Juks/plagiat: Eksamen skal være et individuelt, selvstendig arbeid. Det er tillatt å bruke hjelpemidler. Under eksamen er det ikke tillatt å kommunisere med andre personer om oppgaven eller å distribuere utkast til svar. Slik kommunikasjon er å anse som juks.
Alle besvarelser blir kontrollert for plagiat. *Du kan lese mer om juks og plagiering på eksamen her*.

Varslinger: Hvis det oppstår behov for å gi beskjeder til kandidatene underveis i eksamen (f.eks. ved feil i oppgavesettet), vil dette bli gjort via varslinger i Inspera. Et varsel vil dukke opp som en dialogboks på skjermen i Inspera. Du kan finne igjen varselet ved å klikke på bjella øverst i høyre hjørne på skjermen. Det vil i tillegg bli sendt SMS til alle kandidater for å sikre at ingen går glipp av viktig informasjon. Ha mobiltelefonen din tilgjengelig.

Vekting av oppgavene: Det er totalt 100 poeng. Oppgaver 1--6 rettes automatisk og teller 5 poeng hver. Oppgaver 7a, 7b, 7c, 7d, 8a, 8b og 9a rettes manuelt og teller 10 poeng hver.

Generell informasjon: I oppgave 1--6 skal dere kun oppgi korrekt svar og ingen begrunnelse. Bruk flere desimaler i mellomregninger enn i det endelige svaret for å unngå avrundingsfeil. **Følg instruksene for antall desimaler i endelig svar. Hvis ikke kan svaret bli vurdert til feil!** For oppgaver 7a, 7b, 7c, 7d, 8a, 8b og 9a skal alle svar begrunnes og all naturlig mellomregning skal inkluderes. Det må være helt tydelig hvordan man har gått frem for å komme frem til svarene.

OM LEVERING: En del av oppgavene krever at du laster opp filer med din løsning. Du skal laste opp en (og bare en) fil for hver av disse oppgavene. Alle filene må være i pdf-format! **Det anbefales sterkt at man laster opp en løsning straks man er ferdig med en oppgave. Man bør ikke laste opp alle løsningene helt på slutten av eksamenstiden, da dette lett kan føre til tekniske problemer.**

Slik svarer du på oppgavene: Alle oppgaver som *ikke* er av typen filopplasting, skal besvares direkte i Inspera. I Inspera lagres svarene dine automatisk hvert 15. sekund.

NB! Klipp og lim fra andre programmer frarådes, da dette kan medføre at formatering og elementer (bilder, tabeller etc.) vil kunne gå tapt.

Filopplasting: Alle filer må være lastet opp i besvarelsen før eksamenstida går ut. Det er lagt til 30 minutter til ordinær eksamenstid for eventuell digitalisering av håndtegninger og opplasting av filer. (Tilleggstida inngår i gjenstående eksamenstid som vises øverst til venstre på skjermen.)

NB! Det er ditt eget ansvar å påse at du laster opp riktig(e) fil(er). Kontroller filene du har lastet opp ved å klikke “Last ned” når du står i filopplastingsoppgaven. Alle filer kan fjernes og byttes ut så lenge prøven er åpen.

Slik digitaliserer du eventuelle håndtegninger
Slik lagrer du dokumentet ditt som PDF.
Slik fjerner du forfatterinformasjon fra filen(e) du skal levere.

Besvarelsen din leveres automatisk når eksamenstida er ute og prøven stenger, forutsatt at minst én oppgave er besvart. Dette skjer selv om du ikke har klikket «Lever og gå tilbake til Dashboard» på siste side i oppgavesettet. Du kan gjenåpne og redigere besvarelsen din så lenge prøven er åpen. Dersom ingen oppgaver er besvart ved prøveslutt, blir ikke besvarelsen din levert. Dette vil anses som “ikke møtt” til eksamen.

Trekk fra eksamen: Blir du syk under eksamen, eller av andre grunner ønsker å levere blankt/trekke deg, gå til “hamburgermenyen” i øvre høyre hjørne og velg «Lever blankt». Dette kan ikke angres selv om prøven fremdeles er åpen.

Tilgang til besvarelse: Du finner besvarelsen din i Arkiv etter at sluttida for eksamen er passert.

1 1A

Oppgave 1

Et datasett består av sju målinger: 4, 6, 5, 3, 5, 1 og 3. Regn ut følgende størrelser. **Bruk tre desimaler i mellomregninger og oppgi svarene avrundet til en desimal**.

- Gjennomsnitt: ☐.
- Median: ☐.
- Empirisk varians: ☐.

Maks poeng: 5

2 1B

Oppgave 1

Et datasett består av sju målinger: 4, 5, 3, 4, 3, 2 og 1. Regn ut følgende størrelser. **Bruk tre desimaler i mellomregninger og oppgi svarene avrundet til en desimal**.

- Gjennomsnitt: ☐.
- Median: ☐.
- Empirisk varians: ☐.

Maks poeng: 5

3 1C

Oppgave 1

Et datasett består av sju målinger: 2, 6, 4, 6, 5, 3 og 3. Regn ut følgende størrelser. **Bruk tre desimaler i mellomregninger og oppgi svarene avrundet til en desimal**.

- Gjennomsnitt: [ ].
- Median: [ ].
- Empirisk varians: [ ].

Maks poeng: 5

4 1D

Oppgave 1

Et datasett består av sju målinger: 2, 2, 1, 5, 4, 2 og 3. Regn ut følgende størrelser. **Bruk tre desimaler i mellomregninger og oppgi svarene avrundet til en desimal**.

- Gjennomsnitt: [ ].
- Median: [ ].
- Empirisk varians: [ ].

Maks poeng: 5

5 2A

Oppgave 2

Anta at X er en standard normalfordelt stokastisk variabel med sannsynlighetstetthet $f_X(x)$. Vi har to hendelser $A = \{X \leq 3\}$ og $B = \{X \geq 3\}$. Bestem om hvert av de følgende fire utsagnene er sant eller usant.

1. $\mathrm{P}(A \cap B) = f_X(3)$.

- Sant
- Usant

2. $\mathrm{P}(A \cap B) = 0$.

- Sant
- Usant

3. A og B er disjunkte hendelser.

- Sant
- Usant

4. A og B er avhengige hendelser.

- Sant
- Usant

Maks poeng: 5

6 2B

Oppgave 2

Anta at X er en gammafordelt stokastisk variabel med sannsynlighetstetthet $f_X(x)$. Vi har to hendelser $A = \{X \leq 2\}$ og $B = \{X \geq 2\}$. Bestem om hvert av de følgende fire utsagnene er sant eller usant.

1. $\mathrm{P}(A \cap B) = f_X(2)$.

- Sant
- Usant

2. $\mathrm{P}(A \cap B) = 0$.

- Sant
- Usant

3. A og B er disjunkte hendelser.

- Sant
- Usant

4. A og B er avhengige hendelser.

- Sant
- Usant

Maks poeng: 5

7 2C

Oppgave 2

Anta at X er en t-fordelt (med 10 frihetsgrader) stokastisk variabel med sannsynlighetstetthet $f_X(x)$. Vi har to hendelser $A = \{X \leq -1\}$ og $B = \{X \geq -1\}$. Bestem om hvert av de følgende fire utsagnene er sant eller usant.

1. $\mathrm{P}(A \cap B) = f_X(-1)$.

- ○ Sant
- ○ Usant

2. $\mathrm{P}(A \cap B) = 0$.

- ○ Sant
- ○ Usant

3. A og B er disjunkte hendelser.

- ○ Sant
- ○ Usant

4. A og B er avhengige hendelser.

- ○ Sant
- ○ Usant

Maks poeng: 5

8 2D

Oppgave 2

Anta at X er en kji-kvadratfordelt (med 10 frihetsgrader) stokastisk variabel med sannsynlighetstetthet $f_X(x)$. Vi har to hendelser $A = \{X \leq 10\}$ og $B = \{X \geq 10\}$. Bestem om hvert av de følgende fire utsagnene er sant eller usant.

1. $\mathrm{P}(A \cap B) = f_X(10)$.

- Sant
- Usant

2. $\mathrm{P}(A \cap B) = 0$.

- Sant
- Usant

3. A og B er disjunkte hendelser.

- Sant
- Usant

4. A og B er avhengige hendelser.

- Sant
- Usant

Maks poeng: 5

9 3A

Oppgave 3

Vi trekker 5 tilfeldig valgte kort fra en vanlig kortstokk (52 kort, 13 av hver farge). Oppgi svarene under som heltall.

- Hva blir antall mulige utfall (antall mulige ikke-ordnede utvalg)?
- På hvor mange måter kan vi trekke 5 kort som alle er av farge hjerter?

Maks poeng: 5

10 3B

Oppgave 3

Vi trekker 5 tilfeldig valgte kort fra en vanlig kortstokk (52 kort, 13 av hver farge). Oppgi svarene under som heltall.

- Hva blir antall mulige utfall (antall mulige ikke-ordnede utvalg)?
- På hvor mange måter kan vi trekke straight flush (alle kort i samme farge og i rekkefølge, f.eks. kløver 8, kløver 9, kløver 10, kløver knekt, kløver dame)? **Merk at Ess kan være første kort eller siste kort in en straight flush.**

Maks poeng: 5

11 3C

Oppgave 3

Vi trekker 5 tilfeldig valgte kort fra en vanlig kortstokk (52 kort, 13 av hver farge). Oppgi svarene under som heltall.

- Hva blir antall mulige utfall (antall mulige ikke-ordnede utvalg)?
- På hvor mange måter kan vi trekke 5 kort som alle er billedkort?

Maks poeng: 5

12 # 3D

Oppgave 3

Vi trekker 5 tilfeldig valgte kort fra en vanlig kortstokk (52 kort, 13 av hver farge). Oppgi svarene under som heltall.

- Hva blir antall mulige utfall (antall mulige ikke-ordnede utvalg)?
- På hvor mange måter kan vi trekke 5 kort hvorav 4 er Ess?

Maks poeng: 5

13 # 4A

Oppgave 4

La X betegne summen av antall øyne når vi kaster to terninger og la $Z = 10X + 5$. Regn ut følgende sannsynligheter og oppgi svarene avrundet til tre desimaler.

- $P(X < 4) =$
- $P(Z \leq 35) =$
- $P(Z = 25) =$
- $P(Z = 25 | X < 4) =$

Maks poeng: 5

14 4B

Oppgave 4

La X betegne summen av antall øyne når vi kaster to terninger og la $Z = 5X + 10$. Regn ut følgende sannsynligheter og oppgi svarene avrundet til tre desimaler.

- $P(X > 10) =$ ☐
- $P(Z \geq 65) =$ ☐
- $P(Z = 70) =$ ☐
- $P(Z = 70 | X > 10) =$ ☐

Maks poeng: 5

15 4C

Oppgave 4

La X betegne summen av antall øyne når vi kaster to terninger og la $Z = 2X + 10$. Regn ut følgende sannsynligheter og oppgi svarene avrundet til tre desimaler.

- $P(X \leq 4) =$ ☐
- $P(Z < 20) =$ ☐
- $P(Z = 14) =$ ☐
- $P(Z = 14 | X \leq 4) =$ ☐

Maks poeng: 5

16 **4D**

Oppgave 4

La X betegne summen av antall øyne når vi kaster to terninger og la $Z = 10X$. Regn ut følgende sannsynligheter og oppgi svarene avrundet til tre desimaler.

- $P(X \leq 4) =$
- $P(Z < 50) =$
- $P(Z = 20) =$
- $P(Z = 20|X \leq 4) =$

Maks poeng: 5

17 **5A**

Oppgave 5

Anta at X er en normalfordelt stokastisk variabel med forventningsverdi 4.5 og standardavvik 3, og at Y er en normalfordelt stokastisk variabel med forventningsverdi -1.5 og standardavvik 4. Anta i tillegg at X og Y er uavhengige stokastiske variabler. Regn ut de følgende sannsynlighetene og oppgi svarene avrundet til tre desimaler.

- $\mathrm{P}(X > 3) =$.
- $\mathrm{P}(X + Y > 3) =$.
- $\mathrm{P}(X + Y > 3|Y = 3) =$.

Maks poeng: 5

18 5B

Oppgave 5

Anta at X er en normalfordelt stokastisk variabel med forventningsverdi 2.5 og standardavvik 3, og at Y er en normalfordelt stokastisk variabel med forventningsverdi -3 og standardavvik 4. Anta i tillegg at X og Y er uavhengige stokastiske variabler. Regn ut de følgende sannsynlighetene og oppgi svarene avrundet til tre desimaler.

- $\mathrm{P}(X > 3) =$ ☐.
- $\mathrm{P}(X + Y > 3) =$ ☐.
- $\mathrm{P}(X + Y > 3 | Y = 2) =$ ☐.

Maks poeng: 5

19 5C

Oppgave 5

Anta at X er en normalfordelt stokastisk variabel med forventningsverdi 2 og standardavvik 4, og at Y er en normalfordelt stokastisk variabel med forventningsverdi 0 og standardavvik 3. Anta i tillegg at X og Y er uavhengige stokastiske variabler. Regn ut de følgende sannsynlighetene og oppgi svarene avrundet til tre desimaler.

- $\mathrm{P}(X > 3) =$ ☐.
- $\mathrm{P}(X + Y > 3) =$ ☐.
- $\mathrm{P}(X + Y > 3 | Y = 1) =$ ☐.

Maks poeng: 5

20 5D

Oppgave 5

Anta at X er en normalfordelt stokastisk variabel med forventningsverdi -1 og standardavvik 4, og at Y er en normalfordelt stokastisk variabel med forventningsverdi -1 og standardavvik 3. Anta i tillegg at X og Y er uavhengige stokastiske variabler. Regn ut de følgende sannsynlighetene og oppgi svarene avrundet til tre desimaler.

- $\mathrm{P}(X > 3) =$ ☐.
- $\mathrm{P}(X + Y > 3) =$ ☐.
- $\mathrm{P}(X + Y > 3 | Y = 3) =$ ☐.

Maks poeng: 5

21 6A

Oppgave 6

Vi ser på to stokastiske variabler X og Y. Vi kjenner forventningsverdiene og variansene: $\mathrm{E}[X] = 2$, $\mathrm{Var}[X] = 2$, $\mathrm{E}[Y] = -1$ og $\mathrm{Var}[Y] = 1$. I tillegg kjenner vi kovariansen $\mathrm{Cov}[X, Y] = 0.5$. Regn ut størrelsene under og oppgi svarene avrundet til en desimal.

- $\mathrm{E}[X + Y] =$ ☐.
- $\mathrm{Var}[X + Y] =$ ☐.

Maks poeng: 5

22 6B

Oppgave 6

Vi ser på to stokastiske variabler X og Y. Vi kjenner forventningsverdiene og variansene: $\mathrm{E}[X] = -1$, $\mathrm{Var}[X] = 1$, $\mathrm{E}[Y] = 1$ og $\mathrm{Var}[Y] = 1.5$. I tillegg kjenner vi kovariansen $\mathrm{Cov}[X, Y] = 0.25$. Regn ut størrelsene under og oppgi svarene avrundet til en desimal.

- $\mathrm{E}[X + Y] =$ ☐.
- $\mathrm{Var}[X + Y] =$ ☐.

Maks poeng: 5

23 6C

Oppgave 6

Vi ser på to stokastiske variabler X og Y. Vi kjenner forventningsverdiene og variansene: $\mathrm{E}[X] = 2$, $\mathrm{Var}[X] = 2$, $\mathrm{E}[Y] = 1$ og $\mathrm{Var}[Y] = 2$. I tillegg kjenner vi kovariansen $\mathrm{Cov}[X, Y] = 0.5$. Regn ut størrelsene under og oppgi svarene avrundet til en desimal.

- $\mathrm{E}[X + Y] =$ ☐.
- $\mathrm{Var}[X + Y] =$ ☐.

Maks poeng: 5

24 6D

Oppgave 6

Vi ser på to stokastiske variabler X og Y. Vi kjenner forventningsverdiene og variansene: $\mathrm{E}[X] = 3$, $\mathrm{Var}[X] = 1.5$, $\mathrm{E}[Y] = -4$ og $\mathrm{Var}[Y] = 1.5$. I tillegg kjenner vi kovariansen $\mathrm{Cov}[X, Y] = 0.25$. Regn ut størrelsene under og oppgi svarene avrundet til en desimal.

- $\mathrm{E}[X + Y] =$ ☐.
- $\mathrm{Var}[X + Y] =$ ☐.

Maks poeng: 5

25 # 7A

Oppgave 7

Anta at lønnsinntekten Y til en tilfeldig valgt lønnstaker fra en populasjon er en kontinuerlig stokastisk variabel med sannsynlighetstettheten

$$f_Y(y) = \begin{cases} \frac{\lambda y_0^\lambda}{y^{\lambda+1}}, & \text{for} \quad y \geq y_0, \\ 0, & \text{for} \quad y < y_0, \end{cases}$$

hvor $\lambda > 0$ er en ukjent parameter og $y_0 = 300\,000$ kr er en kjent minstelønn fastsatt av myndighetene.

a)

Anta i dette punktet at $\lambda = 1.0$.

- Finn den kumulative fordelingsfunksjonen til lønnsinntekt Y.
- Bruk den kumulative fordelingsfunksjonen til å beregne sannsynligheten for at den laveste lønnsinntekten i et tilfeldig utvalg på $n = 10$ lønnstakere er mindre enn 310 000 kroner.

Vi observerer et tilfeldig utvalg av $n = 10$ lønnsinntekter, $y_1, y_2, \ldots, y_n$, gitt ved 456000, 578000, 325000, 324000, 382000, 1498000, 594000, 405000, 510000 og 326000 kr. Det oppgis at $\sum_{i=1}^{n} \ln(y_i/y_0) = 4.68$.

b)

- Utled sannsynlighetsmaksimeringsestimatoren av λ for et tilfeldig utvalg av størrelse n.
- Beregn sannsynlighetsmaksimeringsestimatet av λ for dataene gitt over.

c)

- La $X = 2\lambda \ln(Y/y_0)$. Vis at X er kji-kvadratfordelt med 2 frihetsgrader. Hva er fordelingen til $\sum_{i=1}^{n} 2\lambda \ln(Y_i/y_0)$? Begrunn svaret.
- Utled et 95%-konfidensintervall for λ basert på et tilfeldig utvalg av størrelse n. Beregn intervallet for dataene oppgitt over.

d)

I denne oppgaven kan du anta kjent at medianen til Y er gitt ved uttrykket $\mu^* = 2^{1/\lambda} y_0$ for $\lambda > 0$. Vi ønsker å teste om dataene oppgitt over gir grunnlag for å si at medianlønnen i populasjonen, μ^*, er større enn 600 000 kr.

- Vis at dette er ekvivalent med hypotesetesten $H_0 : \lambda = 1.0$ versus $H_1 : \lambda < 1.0$.
- Konstruer og gjennomfør testen for dataene over og med et signifikansnivå lik 0.05. Hva blir konklusjonen?

Last opp filen her. Maks én fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 40

26 7B

Oppgave 7

Anta at lønnsinntekten Y til en tilfeldig valgt lønnstaker fra en populasjon er en kontinuerlig stokastisk variabel med sannsynlighetstettheten

$$f_Y(y) = \begin{cases} \frac{\lambda y_0^\lambda}{y^{\lambda+1}}, & \text{for} \quad y \geq y_0, \\ 0, & \text{for} \quad y < y_0, \end{cases}$$

hvor $\lambda > 0$ er en ukjent parameter og $y_0 = 300\,000$ kr er en kjent minstelønn fastsatt av myndighetene.

a)

Anta i dette punktet at $\lambda = 1.36$.

- Finn den kumulative fordelingsfunksjonen til lønnsinntekt Y.
- Bruk den kumulative fordelingsfunksjonen til å beregne sannsynligheten for at den laveste lønnsinntekten i et tilfeldig utvalg på $n = 8$ lønnstakere er mindre enn 310 000 kroner.

Vi observerer et tilfeldig utvalg av $n = 8$ lønnsinntekter, $y_1, y_2, \ldots, y_n$, gitt ved 417000, 501000, 320000, 319000, 363000, 1056000, 512000, 379000 kr. Det oppgis at $\sum_{i=1}^n \ln(y_i/y_0) = 3.19$.

b)

- Utled sannsynlighetsmaksimeringsestimatoren av λ for et tilfeldig utvalg av størrelse n.
- Beregn sannsynlighetsmaksimeringsestimatet av λ for dataene gitt over.

c)

- La $X = 2\lambda \ln(Y/y_0)$. Vis at X er kji-kvadratfordelt med 2 frihetsgrader. Hva er fordelingen til $\sum_{i=1}^n 2\lambda \ln(Y_i/y_0)$? Begrunn svaret.
- Utled et 95%-konfidensintervall for λ basert på et tilfeldig utvalg av størrelse n. Beregn intervallet for dataene oppgitt over.

d)

I denne oppgaven kan du anta kjent at medianen til Y er gitt ved uttrykket $\mu^* = 2^{1/\lambda} y_0$ for $\lambda > 0$. Vi ønsker å teste om dataene oppgitt over gir grunnlag for å si at medianlønnen i populasjonen, μ^*, er større enn 500 000 kr.

- Vis at dette er ekvivalent med hypotesetesten $H_0 : \lambda = 1.36$ versus $H_1 : \lambda < 1.36$.
- Konstruer og gjennomfør testen for dataene over og med et signifikansnivå lik 0.05. Hva blir konklusjonen?

Last opp filen her. Maks én fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 40

27 **7C**

Oppgave 7

Anta at lønnsinntekten Y til en tilfeldig valgt lønnstaker fra en populasjon er en kontinuerlig stokastisk variabel med sannsynlighetstettheten

$$f_Y(y) = \begin{cases} \frac{\lambda y_0^\lambda}{y^{\lambda+1}}, & \text{for} \quad y \geq y_0, \\ 0, & \text{for} \quad y < y_0, \end{cases}$$

hvor $\lambda > 0$ er en ukjent parameter og $y_0 = 400\,000$ kr er en kjent minstelønn fastsatt av myndighetene.

a)

Anta i dette punktet at $\lambda = 1.24$.

- Finn den kumulative fordelingsfunksjonen til lønnsinntekt Y.
- Bruk den kumulative fordelingsfunksjonen til å beregne sannsynligheten for at den laveste lønnsinntekten i et tilfeldig utvalg på $n = 6$ lønnstakere er mindre enn 410 000 kroner.

Vi observerer et tilfeldig utvalg av $n = 6$ lønnsinntekter, $y_1, y_2, \ldots, y_n$, gitt ved 514000, 593000, 420000, 419000, 463000, 1050000 kr. Det oppgis at $\sum_{i=1}^{n} \ln(y_i/y_0) = 1.85$.

b)

- Utled sannsynlighetsmaksimeringsestimatoren av λ for et tilfeldig utvalg av størrelse n.
- Beregn sannsynlighetsmaksimeringsestimatet av λ for dataene gitt over.

c)

- La $X = 2\lambda \ln(Y/y_0)$. Vis at X er kji-kvadratfordelt med 2 frihetsgrader. Hva er fordelingen til $\sum_{i=1}^{n} 2\lambda \ln(Y_i/y_0)$? Begrunn svaret.
- Utled et 95%-konfidensintervall for λ basert på et tilfeldig utvalg av størrelse n. Beregn intervallet for dataene oppgitt over.

d)

I denne oppgaven kan du anta kjent at medianen til Y er gitt ved uttrykket $\mu^* = 2^{1/\lambda} y_0$ for $\lambda > 0$. Vi ønsker å teste om dataene oppgitt over gir grunnlag for å si at medianlønnen i populasjonen, μ^*, er større enn 700 000 kr.

- Vis at dette er ekvivalent med hypotesetesten $H_0 : \lambda = 1.24$ versus $H_1 : \lambda < 1.24$.
- Konstruer og gjennomfør testen for dataene over og med et signifikansnivå lik 0.05. Hva blir konklusjonen?

Last opp filen her. Maks én fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 40

28 7D

Oppgave 7

Anta at lønnsinntekten Y til en tilfeldig valgt lønnstaker fra en populasjon er en kontinuerlig stokastisk variabel med sannsynlighetstettheten

$$f_Y(y) = \begin{cases} \frac{\lambda y_0^\lambda}{y^{\lambda+1}}, & \text{for} \quad y \geq y_0, \\ 0, & \text{for} \quad y < y_0, \end{cases}$$

hvor $\lambda > 0$ er en ukjent parameter og $y_0 = 400\,000$ kr er en kjent minstelønn fastsatt av myndighetene.

a)

Anta i dette punktet at $\lambda = 3.11$.

- Finn den kumulative fordelingsfunksjonen til lønnsinntekt Y.
- Bruk den kumulative fordelingsfunksjonen til å beregne sannsynligheten for at den laveste lønnsinntekten i et tilfeldig utvalg på $n = 12$ lønnstakere er mindre enn 410 000 kroner.

Vi observerer et tilfeldig utvalg av $n = 12$ lønnsinntekter, $y_1, y_2, \ldots, y_n$, gitt ved 584000, 722000, 430000, 429000, 497000, 1701000, 740000, 524000, 645000, 431000, 802000, 586000 kr. Det oppgis at $\sum_{i=1}^n \ln(y_i/y_0) = 5.29$.

b)

- Utled sannsynlighetsmaksimeringsestimatoren av λ for et tilfeldig utvalg av størrelse n.
- Beregn sannsynlighetsmaksimeringsestimatet av λ for dataene gitt over.

c)

- La $X = 2\lambda \ln(Y/y_0)$. Vis at X er kji-kvadratfordelt med 2 frihetsgrader. Hva er fordelingen til $\sum_{i=1}^n 2\lambda \ln(Y_i/y_0)$? Begrunn svaret.
- Utled et 95%-konfidensintervall for λ basert på et tilfeldig utvalg av størrelse n. Beregn intervallet for dataene oppgitt over.

d)

I denne oppgaven kan du anta kjent at medianen til Y er gitt ved uttrykket $\mu^* = 2^{1/\lambda} y_0$ for $\lambda > 0$. Vi ønsker å teste om dataene oppgitt over gir grunnlag for å si at medianlønnen i populasjonen, μ^*, er større enn 500 000 kr.

- Vis at dette er ekvivalent med hypotesetesten $H_0 : \lambda = 3.11$ versus $H_1 : \lambda < 3.11$.
- Konstruer og gjennomfør testen for dataene over og med et signifikansnivå lik 0.05. Hva blir konklusjonen?

Last opp filen her. Maks én fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 40

29 8A

Oppgave 8

Vi ser på en normalpopulasjon med ukjent forventningsverdi μ og kjent varians $\sigma^2 = 1$. Målet er å teste nullhypotesen $H_0 : \mu = 5$ mot alternativ hypotese $H_1 : \mu < 5$. Dette skal gjøres ved signifikansnivå $\alpha = 0.05$ med et tilfeldig utvalg $X_1, X_2, \ldots, X_n$ fra normalpopulasjonen.

a)

Anta at man velger å forkaste H_0 dersom $\bar{x} \leq 5 - \frac{z_\alpha}{\sqrt{n}}$, der $\bar{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$ og z_α betegner verdien slik at $\mathrm{P}(Z > z_\alpha) = \alpha$ for en standardnormalfordelt stokastisk variabel Z.

- Regn ut sannsynligheten for å gjøre en type I feil.
- Regn ut styrken til hypotesetesten når $n = 10$ og sann forventningsverdi er $\mu = 4.5$.

b)

Anta at en forsker gjør 2 uavhengige tilfeldige utvalg hver av størrelse $n = 10$. Anta nå at man i stedet for beslutningsregelen angitt i a) velger å forkaste H_0 dersom minst ett av de to utvalgene oppfyller forkastningsregelen $\bar{x} \leq 5 - \frac{z_\alpha}{\sqrt{n}}$.

- Regn ut sannsynligheten for å gjøre en type I feil med forskerens fremgangsmåte. Det vil si sannsynligheten for at minst ett av de to utvalgene oppfyller forkastningsregelen i a) når nullhypotesen er sann.
- Forklar hvorfor sannsynligheten for en type I feil er større i forskerens framgangsmåte enn svaret i a).
- Vi skal nå endre beslutningsregelen og forkaste H_0 dersom minst ett av de to utvalgene oppfyller forkastningsregelen $\bar{x} \leq k$. Bestem tallverdien til k som gir signifikansnivå 0.05.

Last opp din fil her. Maksimum en fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 20

30 # 8B

Oppgave 8

Vi ser på en normalpopulasjon med ukjent forventningsverdi μ og kjent varians $\sigma^2 = 1$. Målet er å teste nullhypotesen $H_0 : \mu = 5$ mot alternativ hypotese $H_1 : \mu < 5$. Dette skal gjøres ved signifikansnivå $\alpha = 0.10$ med et tilfeldig utvalg $X_1, X_2, \ldots, X_n$ fra normalpopulasjonen.

a)

Anta at man velger å forkaste H_0 dersom $\bar{x} \leq 5 - \frac{z_\alpha}{\sqrt{n}}$, der $\bar{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$ og z_α betegner verdien slik at $\mathrm{P}(Z > z_\alpha) = \alpha$ for en standardnormalfordelt stokastisk variabel Z.

- Regn ut sannsynligheten for å gjøre en type I feil.
- Regn ut styrken til hypotesetesten når $n = 10$ og sann forventningsverdi er $\mu = 4.5$.

b)

Anta at en forsker gjør 2 uavhengige tilfeldige utvalg hver av størrelse $n = 10$. Anta nå at man i stedet for beslutningsregelen angitt i a) velger å forkaste H_0 dersom minst ett av de to utvalgene oppfyller forkastningsregelen $\bar{x} \leq 5 - \frac{z_\alpha}{\sqrt{n}}$.

- Regn ut sannsynligheten for å gjøre en type I feil med forskerens fremgangsmåte. Det vil si sannsynligheten for at minst ett av de to utvalgene oppfyller forkastningsregelen i a) når nullhypotesen er sann.
- Forklar hvorfor sannsynligheten for en type I feil er større i forskerens framgangsmåte enn svaret i a).
- Vi skal nå endre beslutningsregelen og forkaste H_0 dersom minst ett av de to utvalgene oppfyller forkastningsregelen $\bar{x} \leq k$. Bestem tallverdien til k som gir signifikansnivå 0.10.

Last opp din fil her. Maksimum en fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 20

31 **8C**

Oppgave 8

Vi ser på en normalpopulasjon med ukjent forventningsverdi μ og kjent varians $\sigma^2 = 1$. Målet er å teste nullhypotesen $H_0 : \mu = 5$ mot alternativ hypotese $H_1 : \mu < 5$. Dette skal gjøres ved signifikansnivå $\alpha = 0.05$ med et tilfeldig utvalg $X_1, X_2, \ldots, X_n$ fra normalpopulasjonen.

a)
Anta at man velger å forkaste H_0 dersom $\bar{x} \leq 5 - \frac{z_\alpha}{\sqrt{n}}$, der $\bar{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$ og z_α betegner verdien slik at $\mathrm{P}(Z > z_\alpha) = \alpha$ for en standardnormalfordelt stokastisk variabel Z.

- Regn ut sannsynligheten for å gjøre en type I feil.
- Regn ut styrken til hypotesetesten når $n = 20$ og sann forventningsverdi er $\mu = 4.5$.

b)
Anta at en forsker gjør 2 uavhengige tilfeldige utvalg hver av størrelse $n = 20$. Anta nå at man i stedet for beslutningsregelen angitt i a) velger å forkaste H_0 dersom minst ett av de to utvalgene oppfyller forkastningsregelen $\bar{x} \leq 5 - \frac{z_\alpha}{\sqrt{n}}$.

- Regn ut sannsynligheten for å gjøre en type I feil med forskerens fremgangsmåte. Det vil si sannsynligheten for at minst ett av de to utvalgene oppfyller forkastningsregelen i a) når nullhypotesen er sann.
- Forklar hvorfor sannsynligheten for en type I feil er større i forskerens framgangsmåte enn svaret i a).
- Vi skal nå endre beslutningsregelen og forkaste H_0 dersom minst ett av de to utvalgene oppfyller forkastningsregelen $\bar{x} \leq k$. Bestem tallverdien til k som gir signifikansnivå 0.05.

Last opp din fil her. Maksimum en fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 20

32 8D

Oppgave 8

Vi ser på en normalpopulasjon med ukjent forventningsverdi μ og kjent varians $\sigma^2 = 1$. Målet er å teste nullhypotesen $H_0 : \mu = 5$ mot alternativ hypotese $H_1 : \mu < 5$. Dette skal gjøres ved signifikansnivå $\alpha = 0.10$ med et tilfeldig utvalg $X_1, X_2, \ldots, X_n$ fra normalpopulasjonen.

a)
Anta at man velger å forkaste H_0 dersom $\bar{x} \leq 5 - \frac{z_\alpha}{\sqrt{n}}$, der $\bar{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$ og z_α betegner verdien slik at $\mathrm{P}(Z > z_\alpha) = \alpha$ for en standardnormalfordelt stokastisk variabel Z.

- Regn ut sannsynligheten for å gjøre en type I feil.
- Regn ut styrken til hypotesetesten når $n = 20$ og sann forventningsverdi er $\mu = 4.5$.

b)
Anta at en forsker gjør 2 uavhengige tilfeldige utvalg hver av størrelse $n = 20$. Anta nå at man i stedet for beslutningsregelen angitt i a) velger å forkaste H_0 dersom minst ett av de to utvalgene oppfyller forkastningsregelen $\bar{x} \leq 5 - \frac{z_\alpha}{\sqrt{n}}$.

- Regn ut sannsynligheten for å gjøre en type I feil med forskerens fremgangsmåte. Det vil si sannsynligheten for at minst ett av de to utvalgene oppfyller forkastningsregelen i a) når nullhypotesen er sann.
- Forklar hvorfor sannsynligheten for en type I feil er større i forskerens framgangsmåte enn svaret i a).
- Vi skal nå endre beslutningsregelen og forkaste H_0 dersom minst ett av de to utvalgene oppfyller forkastningsregelen $\bar{x} \leq k$. Bestem tallverdien til k som gir signifikansnivå 0.10.

Last opp din fil her. Maksimum en fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 20

33 # 9A

Oppgave 9

Vi bruker den lineære regresjonsmodellen

$$Y_i = \ln(x_i)\beta + \epsilon_i, \quad i = 1, 2, \ldots, n,$$

hvor $\ln$ betegner den naturlige logaritmen, $x_i > 0$ er en kovariat, β er koeffisienten til kovariaten, og $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ er uavhengige og normalfordelte stokastiske variabler med samme forventningsverdi 0 og **kjente** varians $\sigma^2 = 2$. Du skal bruke estimatoren $\hat{\beta} = \frac{\sum_{i=1}^n \ln(x_i) Y_i}{\sum_{i=1}^n (\ln(x_i))^2}$ for β.

- Vis at $\hat{\beta}$ er forventningsrett og bestem variansen til $\hat{\beta}$.
- Skriv opp et uttrykk for et 95% prediksjonsintervall for en ny verdi Y med kovariatverdi x. Du må begrunne valg av uttrykk, men trenger ikke å utlede det.
- Finn tallverdien til prediksjonsintervallet når $x = 2$, $n = 10$, $\sum_{i=1}^{10} \ln(x_i) y_i = 4$ og $\sum_{i=1}^{10} (\ln(x_i))^2 = 9.3$.

Last opp din fil her. Maksimum en fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 10

34 9B

Oppgave 9

Vi bruker den lineære regresjonsmodellen

$$Y_i = x_i^3\beta + \epsilon_i, \quad i = 1, 2, \ldots, n,$$

hvor x_i er en kovariat, β er koeffisienten til kovariaten, og $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ er uavhengige og normalfordelte stokastiske variabler med samme forventningsverdi 0 og **kjente** varians $\sigma^2 = 2$. Du skal bruke estimatoren $\hat{\beta} = \frac{\sum_{i=1}^n x_i^3 Y_i}{\sum_{i=1}^n x_i^6}$ for β.

- Vis at $\hat{\beta}$ er forventningsrett og bestem variansen til $\hat{\beta}$.
- Skriv opp et uttrykk for et 95% prediksjonsintervall for en ny verdi Y med kovariatverdi x. Du må begrunne valg av uttrykk, men trenger ikke å utlede det.
- Finn tallverdien til prediksjonsintervallet når $x = 2$, $n = 10$, $\sum_{i=1}^{10} x_i^3 y_i = 3$ og $\sum_{i=1}^{10} x_i^6 = 10$.

Last opp din fil her. Maksimum en fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 10

35 **9C**

Oppgave 9

Vi bruker den lineære regresjonsmodellen

$$Y_i = \mathrm{e}^{x_i}\beta + \epsilon_i, \quad i = 1, 2, \ldots, n,$$

hvor x_i er en kovariat, β er koeffisienten til kovariaten, og $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ er uavhengige og normalfordelte stokastiske variabler med samme forventningsverdi 0 og **kjente** varians $\sigma^2 = 2$. Du skal bruke estimatoren $\hat{\beta} = \frac{\sum_{i=1}^{n} \mathrm{e}^{x_i} Y_i}{\sum_{i=1}^{n} \mathrm{e}^{2x_i}}$ for β.

- Vis at $\hat{\beta}$ er forventningsrett og bestem variansen til $\hat{\beta}$.
- Skriv opp et uttrykk for et 95% prediksjonsintervall for en ny verdi Y med kovariatverdi x. Du må begrunne valg av uttrykk, men trenger ikke å utlede det.
- Finn tallverdien til prediksjonsintervallet når $x = 2$, $n = 10$, $\sum_{i=1}^{10} \mathrm{e}^{x_i} y_i = 5$ og $\sum_{i=1}^{10} \mathrm{e}^{2x_i} = 15$.

Last opp din fil her. Maksimum en fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 10

36 # 9D

Oppgave 9

Vi bruker den lineære regresjonsmodellen

$$Y_i = \sqrt{x_i}\beta + \epsilon_i, \quad i = 1, 2, \dots, n,$$

hvor $x_i > 0$ er en kovariat, β er koeffisienten til kovariaten, og $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ er uavhengige og normalfordelte stokastiske variabler med samme forventningsverdi 0 og **kjente** varians $\sigma^2 = 2$.
Du skal bruke estimatoren $\hat{\beta} = \frac{\sum_{i=1}^n \sqrt{x_i} Y_i}{\sum_{i=1}^n x_i}$ for β.

- Vis at $\hat{\beta}$ er forventningsrett og bestem variansen til $\hat{\beta}$.
- Skriv opp et uttrykk for et 95% prediksjonsintervall for en ny verdi Y med kovariatverdi x. Du må begrunne valg av uttrykk, men trenger ikke å utlede det.
- Finn tallverdien til prediksjonsintervallet når $x = 2$, $n = 10$, $\sum_{i=1}^{10} \sqrt{x_i} y_i = 100$ og $\sum_{i=1}^{10} x_i = 50$.

Last opp din fil her. Maksimum en fil.

Følgende filtyper er tillatt: **.pdf** Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 10