

i Institutt for matematiske fag

Eksamensoppgave i ISTA1001, ISTA1002, ISTA1003, ISTG1001, ISTG1002, ISTG1003, ISTT1001, ISTT1002, ISTT1003 Statistikk

Eksamensdato: PRØVEEKSAMEN oktober 2021

Eksamenstid (fra-til): 09:00 – 12:00

Hjelpemiddelkode/Tillatte hjelpemidler: C

Godkjent kalkulator

Digitale ressurser under eksamen: Formelark og fem tabeller (binomisk kumulativfordeling, Poisson kumulativfordeling, normal kumulativfordeling, normalfordeling kritisk verdi og t-fordeling kritisk verdi) er lagt ved eksamen som pdf-filer

Faglig kontakt under eksamen: Mette Langaas (988 476 49)

ANNEN INFORMASJON:

Henvend deg til en eksamensvakt hvis du ønsker å kontakte faglig kontaktperson under eksamen. Noter spørsmålet ditt på forhånd. Faglig kontaktperson skal kun kontaktes dersom det er direkte feil eller mangler i oppgavesettet.

Tallsvar: For hver oppgave der du skal skrive inn et tallsvar er det oppgitt antall desimaler du skal skrive inn. Både desimalkomma og desimalpunktum kan brukes.

Vekting av oppgavene: er oppgitt for hver oppgave.

Lagring: Besvarelsen din i Inspira Assessment lagres automatisk hvert 15. sekund.

Varslinger: Hvis det oppstår behov for å gi beskjeder til kandidatene underveis i eksamen (f.eks. ved klare mangler i oppgavesettet), vil dette bli gjort via varslinger i Inspira. Et varsel vil dukke opp som en dialogboks på skjermen i Inspira. Du kan finne igjen varselet ved å klikke på bjella øverst i høyre hjørne på skjermen.

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. Tilkall så eksamensvakt og følg instruksene fra eksamensvakten.

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

1 Oppgave 1: Hendingar

Innledning:

La oss anta at en tilfeldig valgt student som jobber med et spørsmål på en eksamen har en sannsynlighet på 0.75 for å kunne teorien og klare utregningene, som gir korrekt svar på spørsmålet (dvs. studenten trenger ikke å tippe).

Hvis studenten ikke finner frem til korrekt svar ved teori og utregning, så har studenten mulighet til å tippe. Hvis studenten tipper, så antar vi at sannsynligheten for at studenten velger det riktige svaret er 0.2 .

Spørsmål:

a) Hva er sannsynligheten for at en tilfeldig valgt student svarer rett på eksamensspørsmålet?

Velg ett alternativ

- ☐ 0.05
- ☐ 0.20
- ☐ 0.50
- ☐ 0.75
- ☐ 0.80

b) Gitt at studenten svarer korrekt på dette spørsmålet, hva er sannsynligheten for at studenten *ikke* tippet? Oppgi svaret med 4 desimaler, f.eks. 0.1234 eller 0.9876.

Maks poeng: 10

2 Oppgave 2: Geometrisk fordelt stokastisk variabel

Innledning: La X være geometrisk fordelt med parameter $p = 0.1$. Det vil si at X har punktsannsynlighet

$$P(X = x) = p(1 - p)^{x-1} \text{ for } x = 1, 2, \dots$$

Spørsmål: Fyll inn riktige verdier for følgende tre sannsynligheter. Oppgi svarene med 2 desimalar, f.eks. 0.12 eller 0.98.

a)

$$P(X = 5) = \boxed{}$$

b)

$$P(X \geq 20) = \boxed{}$$

c)

$$P(X < 5 | X \geq 3) = \boxed{}$$

Maks poeng: 15

3 Oppgave 3: Normalfordelte rør

Innledning:

En bedrift produserer rør som settes sammen til gassrørledninger. Vi antar at lengden til et rør er normalfordelt med forventningsverdi 12 meter og standardavvik 0.1 meter. Vi antar videre at lengden til rørene som produseres er uavhengige av hverandre.

Spørsmål:

a) Hva er sannsynligheten for at et tilfeldig valgt rør er lengre enn 12.2 meter?

Skriv inn svaret med 4 desimalar, f.eks. 0.1234 eller 0.9876.

b) Dersom standardavviket til lengdene på rørene kunne justeres til en annen verdi, hvilken verdi måtte det justeres til for at sannsynligheten skal være 0.95 for at lengden til et tilfeldig valgt rør er mellom 11.90 og 12.10 meter?

Velg ett alternativ

☐ 0.01

☐ 0.02

☐ 0.05

☐ 0.10

☐ 0.12

☐ 0.15

Innledning:

Røren som brukes er fortsatt normalfordelte med forventningsverdi 12 meter og standardavvik 0.1 meter. Bedriften har fått i oppdrag å produsere en 10800 meter lang rørledning, og planlegger å produsere 900 rør til denne rørledningen. Totallengden på rørledningen blir da summen av lengdene til de 900 rørene.

Spørsmål:

c) Hva er forventningsverdi og varians til totallengden av rørledningen? Skriv inn svarene som heltall, f.eks. 1 eller 100.

Forventningsverdi:

Varians:

d) Hvis rørledningen er kortare enn 10795 meter eller lengre enn 10805 meter vil kunden kunne klage på leveransen fra bedriften. Hvor sannsynlig er det at kunden klager? Skriv inn svaret med

3 desimaler, f.eks. 0.123 eller 0.987.

Maks poeng: 15

4 Oppgave 4: Forventning og varians

Innledning: La X_1 og X_2 være to stokastiske variabler med $E(X_1) = 0$, $E(X_2) = -1$, $\text{Var}(X_1) = 3$, $\text{Var}(X_2) = 2$ og $\text{Cov}(X_1, X_2) = 1$. Anta videre at vi har en stokastisk variabel Y som er uavhengig av X_1 og X_2 og som har $E(Y) = 3$ og $\text{Var}(Y) = 2$.

La stokastiske variabler Z_1 og Z_2 være gitt som

$$Z_1 = 3X_1 + Y \quad \text{og} \quad Z_2 = X_1 - 4X_2.$$

Spørsmål: Finn forventningsverdi og varians for Z_1 og for Z_2 . Angi tallene som heltall, f.eks. 1 eller 100.

$$E(Z_1) = \text{$$

$$\text{Var}(Z_1) = \text{$$

$$E(Z_2) = \text{$$

$$\text{Var}(Z_2) = \text{$$

Maks poeng: 15

5 Oppgave 5: Konfidensintervall

Innledning: Vi ser på en lignende problemstilling som i oppgave 3, der en bedrift produserer rør som settes sammen til gassrørledninger. Vi antar også her at lengden til et rør er normalfordelt, men nå antar vi at forventningsverdien μ er ukjent og at standardavvik er kjent og lik $\sigma = 0.1$ meter. Vi antar fremdeles at lengden til rørene som produseres er uavhengige av hverandre.

Bedriften produserer $n = 12$ rør og måler lengden på disse. Gjennomsnittlig lengde ble 11.97 meter.

Spørsmål:

a) Regn ut et 95% konfidensintervall for μ og oppgi nedre og øvre grense med to desimaler, f.eks. 5.45 eller 32.76.

Nedre grense:

Øvre grense:

b) Hvor mange målinger må man i denne situasjonen minst gjøre dersom man ønsker et 95% konfidensintervall for μ med lengde maksimalt 0.02?

Velg ett alternativ

- ☐ 20
- ☐ 192
- ☐ 193
- ☐ 271
- ☐ 385

Maks poeng: 15

6 Oppgave 6: Hypotesetest for fosfor i utslippsvann

Innledning:

I denne oppgaven skal vi se på utslipp av fosfor fra et kommunalt kloakkrenseanlegg. Renseanlegget har tatt i bruk nytt utstyr og fått oppgitt at mengdene fosfor i like store prøver tatt på ulike dager vil være uavhengige og normalfordelte med forventningsverdi μ mg/l som ikke overskrider 0.15 mg/l, og standardavvik $\sigma = 0.02$ mg/l.

Spørsmål:

a) Vi vil bruke målinger av fosforinnhold i utslippsvannet til å utføre en hypotesetest der vi tester om vi har grunnlag for å konkludere at gjennomsnittlig mengde fosfor i utslippsvannet i det lange løp skal være større enn 0.15 mg/l, slik at tiltak må iverksettes.

Hvilken alternativ hypotese setter vi opp?

Velg alternativ ▼ ($\mu < 0.15$, $\mu \neq 0.15$, $\mu > 0.15$, $\mu = 0.15$)

Innledning:

b) Vi har samlet inn data og får oppgitt at gjennomsnittet av fosforinnholdet i 12 uavhengige prøver av utslippsvannet ble 0.158 mg/l. Bruk signifikansnivå 5% og utfør hypotesetesten. Hva blir konklusjonen?

Velg ett alternativ

- ☐ Testobservatoren er større enn kritisk verdi og vi forkaster ikke nullhypotesen.
- ☐ Testobservatoren er mindre enn kritisk verdi og vi forkaster ikke nullhypotesen.
- ☐ Testobservatoren er mindre enn kritisk verdi og vi forkaster nullhypotesen.
- ☐ Testobservatoren er større enn kritisk verdi og vi forkaster nullhypotesen.

c) Regn ut p -verdien til testen. Oppgi svaret med 4 desimaler, f.eks. 0.1234 eller 0.9876.

d) Ledelsen ved rensesanlegget ønsker å studere testen fra b) nærmere.

Vi antar fremdeles at $\sigma = 0.02$ mg/l og at signifikansnivå 5% benyttes, og at testen er basert på 12 målinger. Regn ut styrken til testen for alternativet $\mu = 0.17$ mg/l. Oppgi svaret med 4 desimaler, f.eks. 0.1234 eller 0.9876.

7 Oppgave 7: Lineær regresjon og blodceller

MERK: Hvis det til eksamen 29.10 blir gitt en oppgave om lineær regresjon, vil følgende oppgave være nivået til en slik oppgave, og mer komplekse aspekter av pensum om lineær regresjon vil **ikke** testes på eksamen 29.10 - men vil være en viktig del av prosjektdelen. Dette er fordi eksamen kommer så tidlig i 2021 (og dette vil da ikke gjelde kontinuasjonseksamen eller senere eksamener i fellesdelen).

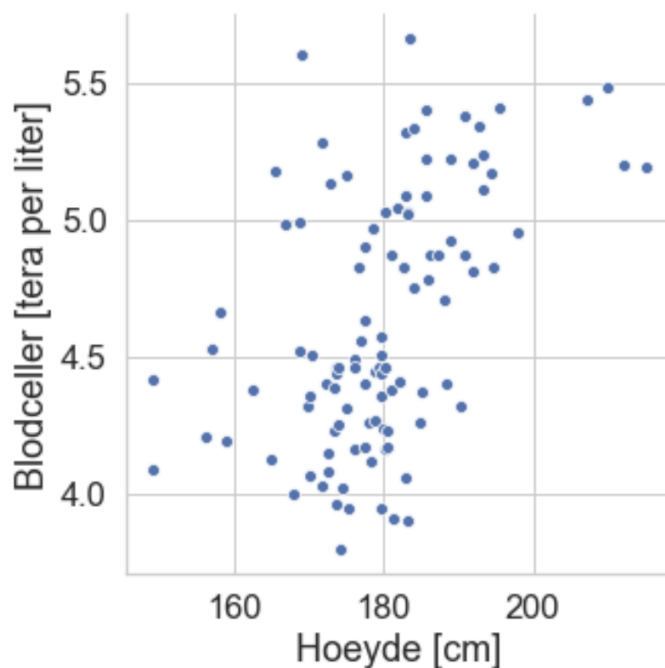
Innledning:

Vi skal studere antallet røde blodceller (per liter blod) til ulike idrettsutøvere. Målet er å undersøke hvordan antallet røde blodceller varierer med høyde.

Datasettet består av informasjon for 105 idrettsutøvere og om inneholder følgende variabler:

- Blodceller: antall røde blodceller, oppgis som antall/ 10^{12} per liter ("tera pr liter" i plottet under)
- Høyde: høyde i cm

Et kryssplott av data er vist under.



Spørsmål:

a) Den empiriske korrelasjonskoeffisienten mellom Høyde og Blodceller kan regnes ut, men du skal fra kryssplottet over velge det korrekte utsagnet om denne empiriske korrelasjonskoeffisienten mellom Høyde og Blodceller:

Velg ett alternativ

- ☐ Den vil være mellom 0 og 1
- ☐ Den vil være mindre enn -1
- ☐ Den vil være større enn 1
- ☐ Den vil være mellom 0 og -1

Innledning: En lineær regresjonsmodell $Y = \beta_0 + \beta_1 x + \varepsilon$ har blitt tilpasset i Python, der responsen Y er Blodceller og forklaringsvariablen (kovariaten) x er Hoeyde. Deler av utskriften (`modell.summary()`) er vist her. Bruk utskriften til å svare på spørsmål b, c og d.

	coef	std err	t	P> t	[0.025	0.975]
Intercept	1.0669	0.630	1.692	0.094	-0.183	2.317
Hoeyde	0.0199	0.004	5.669	0.000	0.013	0.027

Spørsmål:

b) Hva ble den estimerte regresjonslinja?

Velg ett alternativ:

- ☐ $\hat{y} = 0.630 + 0.004 x$
- ☐ $\hat{y} = 1.0669 + 0.630 x$
- ☐ $\hat{y} = 0.0199 + 0.004 x$
- ☐ $\hat{y} = 1.0669 + 0.0199 x$
- ☐ $\hat{y} = 0.630 + 1.692 x$
- ☐ $\hat{y} = 0.0199 + 5.669 x$

c) Hvilken hypotesetest utføres i utskriften i linjen som starter med "Hoeyde" ?

Velg ett alternativ

- ☐ $H_0 : \beta_1 = 1 \text{ mot } H_1 : \beta_1 \neq 1$
- ☐ $H_0 : x \leq y \text{ mot } H_1 : x < y$
- ☐ $H_0 : \beta_1 = 0.0199 \text{ mot } H_1 : \beta_1 \neq 0.0199$
- ☐ $H_0 : \beta_1 = 0 \text{ mot } H_1 : \beta_1 < 0$
- ☐ $H_0 : \beta_1 = 0 \text{ mot } H_1 : \beta_1 > 0$
- ☐ $H_0 : \beta_1 = 0 \text{ mot } H_1 : \beta_1 \neq 0$
- ☐ $H_0 : \beta_1 = 0.004 \text{ mot } H_1 : \beta_1 \neq 0.004$

d) Se på linjen som starter med "Hoeyde". Hvilket av følgende utsagn er korrekte?

Velg ett alternativ:

- ☐ Kovariaten "Hoeyde" har ingen effekt i regresjonen
- ☐ Vi beholder nullhypotesen
- ☐ Vi forkaster nullhypotesen
- ☐ Stigningstallet er ikke forskjellig fra 0

Maks poeng: 15

8 Ekstraoppgave 8: Punktestimering i laboratoriet

MERK: Denne oppgaven er en ekstraoppgave til prøveeksamen oktober 2021 – med mål å gi studentene mer erfaring med eksamensoppgaver innen punktestimering i eksamensoppgaveformatet. Når det gjelder punktestimering så vil det være fokus på å regne forventning og varians av estimatorer, så det vil være kobling til oppgaver der forventning og varians av matematiske uttrykk regnes ut.

Denne oppgave teller 15 poeng og er ikke med i de 100 poengene som oppgave 1-7 som prøveeksamen omfatter.

Innledning:

En laborant skal bestemme konsentrasjonen av et bestemt stoff i en prøve. La μ betegne denne ukjente konsentrasjonen.

Til å måle konsentrasjonen har laboranten tilgjengelig to måleinstrumenter som begge gir normalfordelte målinger med forventningsverdi μ , men de to måleinstrumentene har ulik nøyaktighet.

Instrument A har standardavvik $\sigma_x = 2.4$ og er ikke så dyrt å bruke, så der gjør laboranten 9 uavhengige målinger $X_1, X_2, \dots, X_9$ av den ukjente konsentrasjonen. La $\bar{X} = \frac{1}{9} \sum_{i=1}^9 X_i$.

Instrument B er dyrere i drift, og har standardavvik $\sigma_y = 0.6$. Laboranten får ikke lov til ta mer enn én måling, Y , med instrument B. De 9 målingene med instrument A er uavhengige av den ene målingen med instrument B.

Oppgaven til laboranten er å finne et godt estimat for μ . Laboranten vurderer to estimatorene:

$$\hat{\mu}_1 = \frac{1}{2}(\bar{X} + Y) \quad \text{og} \quad \hat{\mu}_2 = 0.36\bar{X} + 0.64Y$$

Spørsmål:

a) Finn forventningsverdien til hver av de to estimatorene.

Velg ett alternativ

- ☐ $E(\hat{\mu}_1) = \mu$ og $E(\hat{\mu}_2) = \mu$
- ☐ $E(\hat{\mu}_1) = \mu\sigma_x^2$ og $E(\hat{\mu}_2) = \mu\sigma_y^2$
- ☐ $E(\hat{\mu}_1) = \mu$ og $E(\hat{\mu}_2) = 0.36\mu + 0.64\sigma_y$
- ☐ $E(\hat{\mu}_1) = \frac{1}{2}\mu$ og $E(\hat{\mu}_2) = \mu$

b) Regn ut numerisk verdi for variansen til $\hat{\mu}_1$, når det oppgis at $\bar{x} = 67.2$ og $y = 66.4$. Oppgi svaret med 2 desimaler, f.eks. 1.23 eller 34.56.

c) Det oppgis at $\text{Var}(\hat{\mu}_2) = 0.23$. Hvilken av de to estimatorene vil du da foretrekke?

Velg ett alternativ

- ☐ $\hat{\mu}_2$
- ☐ $\hat{\mu}_1$
- ☐ De er like gode

d) Anta at laboranten velger å bruke estimatoren $\hat{\mu}_2$. Hvilken fordeling har estimatoren?

Velg ett alternativ

- ☐ Poisson-fordelt
- ☐ Weibullfordelt
- ☐ Geometrisk fordelt
- ☐ Normalfordelt
- ☐ Eksponentialfordelt
- ☐ Binomisk fordelt

Maks poeng: 15

9 Ekstraoppgave 9: Levetid og pålitelighet

MERK: Denne oppgaven er en ekstraoppgave til prøveeksamen oktober 2021 – med mål å gi studentene mer erfaring med eksamensoppgaver fra den nye delen av pensum om levetid og pålitelighet.

Denne oppgave teller 15 poeng og er ikke med i de 100 poengene som oppgave 1-7 som prøveeksamen omfatter.

Innledning: Levetiden til en bestemt type LED-lys er eksponentialfordelt med forventning 10 år. La T være levetiden til et tilfeldig LED-lys, og anta at levetiden til ulike LED-lys er uavhengige av hverandre.

Spørsmål:

a) Hva er sannsynligheten for at et LED-lys av denne typen virker lenger enn 1 år. Oppgi svaret med 3 desimaler, f.eks. 0.123 eller 0.987.

b) I et rom installeres 2 LED-lys av den aktuelle typen. Anta at levetiden til de to lysene er uavhengig av hverandre. Finn sannsynligheten for at minst ett LED-lys feiler i løpet av 1 år. Oppgi svaret med 3 desimaler, f.eks. 0.123 eller 0.987.

c) I et annet rom er det 6 LED-lys. Anta at levetiden til de 6 lysene er uavhengige av hverandre. Hva er sannsynligheten for at alle LED-lysene virker i minst 5 år? Oppgi svaret med 3 desimaler, f.eks. 0.123 eller 0.987.

Maks poeng: 15