ℹ Institutt for matematiske fag

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

Eksamensdato: PRØVEEKSAMEN oktober 2021

Eksamenstid (frå-til): 09:00 – 12:00

Hjelpemiddelkode/Tillatne hjelpemiddel: C
Godkjent kalkulator
Digitale ressursar under eksamen: Formelark og fem tabellar (binomisk kumulativfordeling, Poisson kumulativfordeling, normal kumulativfordling, normalfordeling kritisk verdi og t-fordeling kritisk verdi) er lagt ved eksamen som pdf-filer

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

ANNEN INFORMASJON:

Vend deg til ei eksamensvakt om du ynskjer å kontakte fagleg kontaktperson. Notér spørsmålet ditt på førehand. Fagleg kontaktperson skal berre kontaktast dersom det er direkte feil eller manglar i oppgåvesettet.

Tallsvar: For kvar oppgåve der du skal skrive inn eit tallsvar er det oppgjeve antal desimalar du skal skrive inn. Både desimalkomma og desimalpunktum kan brukast.

Vekting av oppgavene: er gjeve for kvar oppgåve.

Lagring: Besvarelsen din i Inspera Assessment lagrast automatisk kvart 15. sekund.

Varslinger: Dersom det oppstår behov for å gje beskjedar til kandidatane medan eksamen er i gang (f.eks. ved feil i oppgåvesettet), vil dette bli gjort via varslingar i Inspera. Eit varsel vil dukke opp som en dialogboks på skjermen i Inspera. Du kan finne att varselet ved å klikke på bjølla i øvre høgre hjørne på skjermen.

Trekk fra eksamen: Blir du sjuk under eksamen, eller av andre grunnar ønskjer å levere blankt/trekke deg, gå til "hamburgermenyen" i øvre høgre hjørne og vel «Lever blankt». Dette kan ikkje angrast sjølv om prøven framleis er open. Tilkall så eksamensvakt og følg instruksane frå eksamensvakta.

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

1 Oppgave 1: Hendingar

Innleiing:
La oss anta at ein tilfeldig vald student som jobbar med eit spørsmål på ein eksamen har eit sannsyn på *0.75* for å kunne teorien og klare utrekningane, som gjer korrekt svar på spørsmålet (dvs. studenten treng ikkje å tippe).

Dersom studenten ikkje finn fram til korrekt svar ved teori og utrekning, har studenten mogelegheit til å tippe. Dersom studenten tippar, så antek vi at sannsynet for at studenten vel det riktige svaret er *0.2*.

Spørsmål:
a) Kva er sannsynet for at ein tilfeldig vald student svarar rett på eksamensspørsmålet?

Vel eitt alternativ

- 0.05
- 0.20
- 0.50
- 0.75
- 0.80

b) Gitt at studenten svarar korrekt på eksamensspørsmålet, hva er sannsynet for at studenten *ikkje* tippa?
Oppgje svaret med 4 desimalar, til dømes 0.1234 eller 0.9876.

Maks poeng: 10

2 Oppgåve 2: Geometrisk fordelt stokastisk variabel

Innleiing: La X vere geometrisk fordelt med parameter $p = 0.1$. Det vil si at X har punktsannsyn

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

Spørsmål: Fyll inn riktige verdiar for følgande tre sannsyn. Oppgje svara med 2 desimalar, til dømes 0.12 eller 0.98.

a)

$P(X = 5)$ =

b)

$P(X \geq 20)$ =

c)

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

Maks poeng: 15

3 Oppgåve 3: Normalfordelte rør

Innleiing:
Ei bedrift produserar rør som setjast saman til gassrørledningar. Vi antek at lengda til eit rør er normalfordelt med forventningsverdi 12 meter og standardavvik 0.1 meter. Vi antek vidare at lengda til røra som produserast er uavhengig av kvarandre.

Spørsmål:
a) Kva er sannsynet for at eit tilfeldig vald rør er lengre enn 12.2 meter?
Skriv inn svaret med 4 desimalar, til dømes 0.1234 eller 0.9876.

b) Dersom standardavviket til lengdene på røra kunne justerast til ein anna verdi,
kva verdi måtte det justerast til for at sannsynet skal vere 0.95 for at lengda til et tilfeldig vald rør er mellom 11.90 og 12.10 meter?

Vel eitt alternativ

- 0.01
- 0.02
- 0.05
- 0.10
- 0.12
- 0.15

Innleiing:
Røra som brukast er fortsatt normalfordelte med forventningsverdi 12 meter og standardavvik 0.1 meter. Bedrifta har fått i oppdrag å produsere ein 10800 meter lang rørledning, og planlegg å produsere 900 rør til denne rørledninga. Totallengda på rørledninga blir då summen av lengdene til dei 900 røra.

Spørsmål:
c) Kva er forventningsverdi og varians til totallengda av rørledninga? Skriv inn svara som heiltall, til dømes 1 eller 100.

Forventningsverdi:

Varians:

d) Viss rørledninga er kortare enn 10795 meter eller lengre enn 10805 meter vil kunden kunne klage på leveransa frå bedriften. Kor sannsynleg er det at kunden klagar? Skriv inn svaret med

3 desimaler, til dømes 0.123 eller 0.987.

Maks poeng: 15

4 Oppgåve 4: Forventning og varians

Innleiing:
La X_1 og X_2 vere to stokastiske variablar med $\mathrm{E}(X_1) = 0, \mathrm{E}(X_2) = -1$, $\mathrm{Var}(X_1) = 3$, $\mathrm{Var}(X_2) = 2$ og $\mathrm{Cov}(X_1, X_2) = 1.$ Anta videre at vi har ein stokastisk variabel Y som er uavhengig av X_1 og X_2 og som har $\mathrm{E}(Y) = 3$ og $\mathrm{Var}(Y) = 2.$

La stokastiske variablar Z_1 og Z_2 vere 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. Angje tala som heiltall, til dømes 1 eller 100.

$\mathrm{E}(Z_1) =$

$\mathrm{Var}(Z_1) =$

$\mathrm{E}(Z_2) =$

$\mathrm{Var}(Z_2) =$

Maks poeng: 15

5 Oppgave 5: Konfidensintervall

Innleiing: Vi ser på ei liknande problemstilling som i oppgåve 3, der ei bedrift produserar rør som setjes samen til gassrørledningar. Vi antek og her at lengda til eit rør er normalfordelt, men no antek vi at forventningsverdien μ er ukjend og at standardavvik er kjend og lik $\sigma = 0.1$ meter. Vi antek fremleis at lengda til røra som blir produsert er uavhengig av kvarandre.

Bedrifta produserar $n = 12$ rør og måler lengda av røra. Gjennomsnittleg lengde vart 11.97 meter.

Spørsmål:

a) Rekn ut eit 95% konfidensintervall for μ og oppgje nedre og øvre grense med 2 desimalar, til dømes 5.45 eller 32.76.

Nedre grense: ____

Øvre grense: ____

b) Kor mange målingar må ein i denne situasjonen minst gjere dersom ein ønsker eit 95% konfidensintervall for μ med lengde maksimalt 0.02?

Vel eitt alternativ

- 20
- 192
- 193
- 271
- 385

Maks poeng: 15

6 Oppgåve 6: Hypotesetest for fosfor i utsleppsvatn

Innleiing:
I denne oppgåva skal vi sjå på utslepp av fosfor frå eit kommunalt kloakkrenseanlegg. Renseanlegget har tatt i bruk nytt utstyr og fått oppgjedd at mengda fosfor i like store prøver tatt på ulike dagar vil vere uavhengige av kvarandre og normalfordelte med forventningsverdi μ mg/l som ikkje overskrid 0.15 mg/l, og standardavvik $\sigma = 0.02$ mg/l.

Spørsmål:
a) Vi vil bruke målinger av fosforinnhald i utsleppsvatnet til å utføre ein hypotesetest der vi testar om vi har grunnlag for å konkludere at gjennomsnittleg mengde fosfor i utsleppsvatnet i det lange løp skal vere større enn 0.15 mg/l, slik at tiltak må iverksetjast.

Kva alternativ hypotese set vi opp?

Vel alternativ (μ = 0.15, μ ≠ 0.15, μ > 0.15, μ < 0.15)

Innleiing:

b) Vi har samla inn data og får oppgjedd at gjennomsnittet av fosforinnhaldet i 12 uavhengige prøver av utsleppsvatnet vart 0.158 mg/l. Bruk signifikansnivå 5% og utfør hypotesetesten.
Kva blir konklusjonen?

Vel eitt alternativ

- Testobservatoren er mindre enn kritisk verdi og vi forkastar nullhypotesen.
- Testobservatoren er større enn kritisk verdi og vi forkastar ikkje nullhypotesen.
- Testobservatoren er større enn kritisk verdi og vi forkastar nullhypotesen.
- Testobservatoren er mindre enn kritisk verdi og vi forkastar ikkje nullhypotesen.

c) Rekn ut *p*-verdien til testen. Oppgje svaret med 4 desimalar, til dømes 0.1234 eller 0.9876.

d) Leiinga ved renseanlegget ønsker å studere testen frå b) nærare.
Vi antak fremleis at $\sigma = 0.02$ mg/l og at signifikansnivå 5% vert nytta, og at testen er basert på 12 målinger. Rekn ut styrken til testen for alternativet $\mu = 0.17$ mg/l. Oppgje svaret med 4 desimalar, til dømes 0.1234 eller 0.9876.

Maks poeng: 15

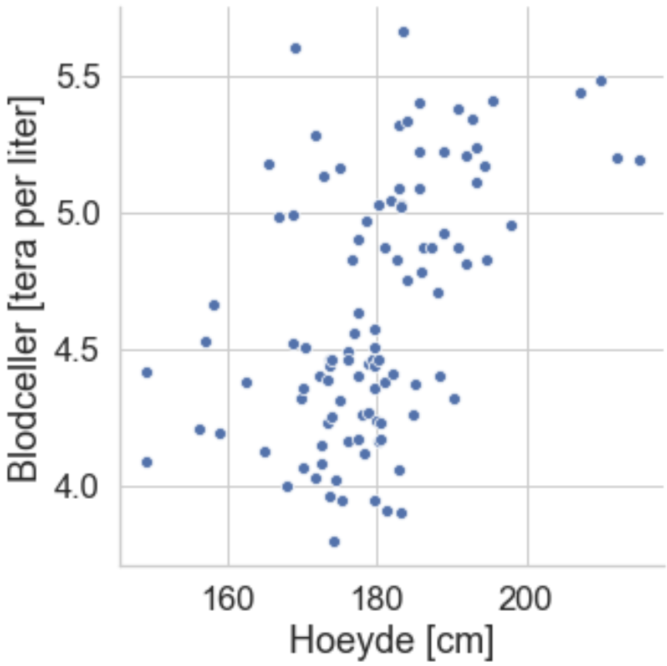
7 Oppgåve 7: Lineær regresjon og blodceller

MERK: Dersom det til eksamen 29.10 vert gjeve ein oppgåve om lineær regresjon, vil følgjande oppgåve være nivået til ein slik oppgåve, og meir komplekse aspekt av pensum om lineær regresjon vil **ikkje** testast på eksamen 29.10 - men vil være ein viktig del av prosjektdelen. Dette er fordi eksamen kjem så tidlig i 2021 (og dette vil da ikkje gjelde kontinuasjonseksamen eller seinare eksamenar i fellesdelen).

Innleiing:
Vi skal studere antalet raue blodceller (per liter blod) til ulike idrettsutøvarar. Målet er å undersøkje om antalet raude blodceller varierar med høgde.
Datasettet består av informasjon for 105 idrettsutøvarar og om inneheld følgjande variablar:

- Blodceller: antal raude blodceller, oppgjedd som antal/10^{12} per liter ("tera pr liter" i plottet under)
- Hoeyde: høgde i cm

Eit kryssplott av data er vist under.



Spørsmål:

a) Den empiriske korrelasjonskoeffisienten mellom Hoeyde og Blodceller kan reknast ut, men du skal frå kryssplottet over velje det korrekte utsagnet om denne empiriske korrelasjonskoeffisienten mellom Hoeyde og Blodceller:

Vel eitt alternativ

- Den vil vere mellom 0 og 1
- Den vil vere mellom 0 og -1
- Den vil vere mindre enn -1
- Den vil vere større enn 1

Innleiing: Ein lineær regresjonsmodell $Y = \beta_0 + \beta_1 x + \varepsilon$ har blitt tilpassa i Python, der responsen Y er Blodceller og forklaringsvariabelen (kovariaten) x er Hoeyde. Deler av utskrifta (*modell.summary()*) er synt her. Bruk utskrifta 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) Kva vart den estimerte regresjonslinja?

Vel eitt alternativ

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

c) Kva hypotesetest utførast i utskrifta i linja som startar med "Hoeyde" ?
Vel eitt alternativ

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

d) Sjå på linja som startar med "Hoeyde". Kva for eit av dei følgjande utsagna er korrekte?
Vel eitt alternativ

- Vi forkastar nullhypotesa
- Kovariaten "Hoeyde" har ingen effekt i regresjonen
- Vi beheld nullhypotesa
- Stigningstallet er ikkje forskjelleg frå 0

Maks poeng: 15

8 Ekstraoppgåve 8: Punktestimering i laboratoriet

MERK: Denne oppgava er ein ekstraoppgåve til prøveeksamen oktober 2021 – med mål å gje studentane meir røynsle med eksamensoppgavar innan punktestimering i eksamensoppgaveformatet. Når det gjeld punktestimering så vil det vere fokus på å rekne forventning og varians av estimatorar, så det vil vere kopling til oppgåver der forventning og varians av matematiske uttrykk reknast ut.
Denne oppgåva teller 15 poeng og er ikkje med i de 100 poenga som oppgave 1-7 som prøveeksamen omfattar.

Innleiing:
Ein laborant skal bestemme konsentrasjonen av eit bestemt stoff i ein prøve. La μ betekne denne ukjende konsentrasjonen.
Til å måle konsentrasjonen har laboranten tilgjengeleg to måleinstrument som begge gjer normalfordelte målingar med forventningsverdi μ , men dei to måleinstrumentane har ulik nøyaktighet.
Instrument A har standardavvik $\sigma_x = 2.4$ og er ikkje så dyrt å bruke, så der gjør laboranten 9 uavhengige målinger $X_1, X_2, \ldots, X_9$ av den ukjente konsentrasjonen. La $\bar{X} = \frac{1}{9}\sum_{i=1}^{9} X_i$.
Instrument B er dyrare i drift, og har standardavvik $\sigma_y = 0.6$. Laboranten får ikkje lov til ta meir enn ei måling, Y, med instrument B. De 9 målingane med instrument A er uavhengige av den eine målinga med instrument B.
Oppgåva til laboranten er å finne eit godt estimat for μ. Laboranten vurderar to estimatorar:
$\hat{\mu}_1 = \frac{1}{2}(\bar{X} + Y) \quad \text{og} \quad \hat{\mu}_2 = 0.36\bar{X} + 0.64Y$

Spørsmål:
a) Finn forventningsverdien til kvar av dei to estimatorane.
Vel eitt alternativ

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

b) Rekn ut numerisk verdi for variansen til $\hat{\mu}_1$, når det blir gitt at $\bar{x} = 67.2$ og $y = 66.4$.
Oppgje svaret med 2 desimalar, til dømes 1.23 eller 34.56.

c) Det blir gitt at $\mathrm{Var}(\hat{\mu}_2) = 0.23$. Kva for ein av dei to estimatorane vil du då foretrekke?

Vel eitt alternativ

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

d) Anta at laboranten vel å bruke estimatoren $\hat{\mu}_2$. Kva fordeling har estimatoren?
Vel eitt alternativ

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

Maks poeng: 15

9 Ekstraoppgåve 9: Levetid og pålitelighet

MERK: Denne oppgava er ein ekstraoppgåve til prøveeksamen oktober 2021 – med mål å gje studentane meir røynsle med eksamensoppgavar frå den nye delen av pensum om levetid og pålitelighet.
Denne oppgåva teller 15 poeng og er ikkje med i de 100 poenga som oppgave 1-7 som prøveeksamen omfattar.

Innleiing: Levetida til ein bestemt type LED-lys er eksponentialfordelt med forventning 10 år. La T vere levetida til eit tilfeldig LED-lys, og anta at levetida til ulike LED-lys er uavhengig av kvarandre.

Spørsmål:

a) Kva er sannsynet for at eit LED-lys av denne typen verkar lenger enn 1 år. Oppgje svaret med 3 desimalar, til dømes 0.123 eller 0.987.

b) I eit rom installerast 2 LED-lys av den aktuelle typen. Anta at levetida til die to lysa er uavhengig av kvarandre. Finn sannsynet for at minst eit LED-lys feilar i løpet av 1 år. Oppgje svaret med 3 desimalar, til dømes 0.123 eller 0.987.

c) I eit anna rom er det 6 LED-lys. Anta at levetiden til dei 6 lysa er uavhengig av kvarandre. Kva er sannsynet for at alle LED-lysa verkar i minst 5 år? Oppgje svaret med 3 desimalar, til dømes 0.123 eller 0.987.

Maks poeng: 15