ℹ Institutt for matematiske fag

Eksamensoppgåve i TMA4240/TMA4245 Statistikk

Eksamensdato: 11. august 2022

Eksamenstid (fra-til): 09.00-13.00

Hjelpemiddelkode/Tillatte hjelpemiddel: Hjelpemiddelkode C.

- Tabellar og formlar i statistikk (Fagbokforlaget).
- Eit gult ark (A5 med stempel) med eigne håndskrevne formlar og notat.
- Bestemt, enkel kalkulator.

Fagleg kontakt under eksamen: Håkon Tjelmeland

Tlf.: 4822 1896

ANNAN INFORMASJON:

Skaff deg overblikk over oppgåvesettet før du byrjar på besvarelsen din.

Generell informasjon:

- I oppgåvene 1 til 4 skal du berre oppgi korrekt svar (altså utan grunngiving) direkte i Inspera. Følg instruksane i oppgåveteksten om tal desimalar i svaret, viss ikkje kan svaret bli vurdert som feil. I alle mellomrekningar må du bruka minst to desimalar meir enn kva du skal ha i svaret.
- I oppgåvene 5 til 7 skal alle svar grunngivast og besvarelsen skal innehalda mellomrekning slik at det er heilt klart korleis ein har tenkt. Besvarelsen på desse oppgåvene skal skrivast for hand på utleverte ark, sjå informasjon under InsperaScan lenger ned på denne sida.

Les oppgåvene nøye, gjer dine eigne antagelser og presiser i besvarelsen kva føresetnader du har lagt til grunn i tolking/avgrensing av oppgåva. Fagleg kontaktperson skal berre kontaktast dersom det er direkte feil eller manglar i oppgåvesettet. Henvend deg til ei eksamensvakt viss du ønsker å kontakta faglærar. Nøter gjerne spørsmålet ditt på førehand.

InsperaScan: Besvarelsen av oppgåve 5, 6 og 7 skal skrivast for hand på utleverte ark. Nedst i kvar oppgåve finn du ein sjusifret kode. Fyll inn denne koden øvst til venstre på arkene du ønsker å levera. Det blir anbefalt å gjera dette undervegs i eksamen. Dersom du treng tilgang til kodane etter at eksamenstida har gått ut, må du klikka «Vis besvarelse».

Vekting av oppgåvene: Vekt ved sensur for kvar deloppgåve er angitt i oppgåvesettet.

Varslingar: Viss det oppstår behov for å gi beskjedar til kandidatane undervegs i eksamen (t.d. ved feil i oppgåvesettet), vil dette bli gjort via varslingar i Inspera. Eit varsel vil dukka opp som ein dialogboks på skjermen. Du kan finna igjen varselet ved å klikka på bjølla øvst til høgre.

Trekk frå/avbroten eksamen: Blir du sjuk under eksamen, eller av andre grunnar ønsker å levera blankt/avbryta eksamen, gå til “hamburgarmenyen” i øvre høgre hjørne og vel «Lever blankt». Dette kan ikkje angrast sjølv om prøven framleis er open.

Tilgang til besvarelse: Etter eksamen finn du besvarelsen din i arkivet i Inspera. Merk at det kan ta éin virkedag før eventuelle handteikningar vil vera tilgjengelege i arkivet.

1 **Innleiing**: La A, B og C vera tre hendingar i eit utfallsrom Følgende sannsyn er oppgitt,

- $P(A) = 0.500,$
- $P(B) = 0.400,$
- $P(C) = 0.300,$
- $P(A \cap B) = 0.300,$
- $P(C|(A \cup B)') = 0.200,$

kvar $(A \cup B)' = S \setminus (A \cup B)$ som vanleg er komplementet til hendinga $A \cup B.$

Oppgåve: Bestem følgjande sannsyn. Angi svara med tre desimalar etter komma.

- $P(A \cup B) =$ ☐
- $P(B|A) =$ ☐
- $P(C|A \cup B) =$ ☐

Maks poeng: 5

2 **Innleiing:** La X og Y vera uavhengige normalfordelte stokastiske variablar. Anta vidare at X har forventningsverdi lik 1 og standardavvik lik 4, mens Y har forventningsverdi lik -2 og standardavvik lik 2.

Oppgåve: Rekn ut følgjande sannsyn. Angi svara med fire desimalar etter komma.

- $P(X > 4) =$
- $P(X + 2Y < 0) =$
- $P(Y - X \leq 4 | Y - X > -3) =$

Maks poeng: 5

3 **Innleiing:** La $X_1, X_2, \ldots, X_n$ vera eit tilfeldig utval frå ei fordeling med forventningsverdi μ og varians σ^2. Basert på dette utvalet er følgjande to stokastiske variablar definert,

$U = \sum_{i=1}^{n} a_i(X_i - \mu)$,

$V = \frac{\sum_{i=1}^{n} a_i X_i}{\sum_{i=1}^{n} a_i}$,

kvar $a_1, a_2, \ldots, a_n$ er konstantar.

Oppgåve: Finn uttrykk for variansane til U og V. Angi under kva som er det korrekte uttrykket for kvar av dei to variansane.

Vel eitt alternativ

- $\mathrm{Var}[U] = (\sigma^2 - \mu)\sum_{i=1}^{n} a_i^2$
- $\mathrm{Var}[U] = (\sigma - \mu)^2\sum_{i=1}^{n} a_i^2$
- $\mathrm{Var}[U] = (\sigma^2 - \mu^2)\sum_{i=1}^{n} a_i^2$
- $\mathrm{Var}[U] = \sigma^2\sum_{i=1}^{n} a_i^2$
- $\mathrm{Var}[U] = \sigma^2\sum_{i=1}^{n} a_i$

Vel eitt alternativ

- $\mathrm{Var}[V] = \frac{\sigma^2}{\sum_{i=1}^{n} a_i}$
- $\mathrm{Var}[V] = \sigma^2$
- $\mathrm{Var}[V] = \sigma^2 \cdot \frac{\sum_{i=1}^{n} a_i^2}{(\sum_{i=1}^{n} a_i)^2}$
- $\mathrm{Var}[V] = \sigma^2 \cdot \frac{(\sum_{i=1}^{n} a_i)^2}{\sum_{i=1}^{n} a_i^2}$
- $\mathrm{Var}[V] = \frac{\sigma^2}{\sum_{i=1}^{n} a_i^2}$

Maks poeng: 5

4 **Innleiing**: La X vera ein kontinuerleg stokastisk variabel med kumulativ fordelingsfunksjon

$$F_X(x) = \begin{cases} 1 - e^{-x^2} & \text{for } x \geq 0, \\ 0 & \text{ellers.} \end{cases}$$

La Y vera gitt ved

$$Y = X^2.$$

Oppgåve: Finn sannsynstettleiken $f_Y(y)$ til Y. Vel under korrekt formel for $f_Y(y)$ for $y \geq 0$.

Vel eitt alternativ

- $f_Y(y) = e^{-y}, y \geq 0$
- $f_Y(y) = 2ye^{-y^2}, y \geq 0$
- $f_Y(y) = 2\sqrt{y}e^{-y}, y \geq 0$
- $f_Y(y) = 4y\sqrt{y}e^{-y}, y \geq 0$
- $f_Y(y) = 1 - e^{-y}, y \geq 0$
- $f_Y(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}, y \geq 0$

Maks poeng: 5

5(a) **Innleiing**: Anta at du kastar to vanlege terningar. La X vera tal seksarar du får, og la Y vera antal terningar der antal auge er tre eller mindre.

Oppgåve: For kvar kombinasjon av $x \in \{0, 1, 2\}$ og $y \in \{0, 1, 2\}$, rekn ut $P(X = x, Y = y)$ og vis at sannsyna blir som oppgitt i følgjande tabell.

$P(X = x, Y = y)$	$x = 0$	$x = 1$	$x = 2$
$y = 0$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{36}$
$y = 1$	$\frac{1}{3}$	$\frac{1}{6}$	0
$y = 2$	$\frac{1}{4}$	0	0

Det blir anbefalt at du løyser denne oppgåva med papir og blyant.

Format

Words: 0

Maks poeng: 10

5(b) **Innleiing**: I dette punktet skal du visa korleis du finn svara med utgangspunkt i sannsyna $P(X = x, Y = y)$ som er angitt i tabellen i punkt 5(a).

Oppgåve: Finn sannsynet $P(X + Y = 2)$.

Bestem marginalt punktsannsyn for X og marginalt punktsannsyn for Y.

Er X og Y uavhengige? Grunngi svaret.

Det blir anbefalt at du løyser denne oppgåva med papir og blyant.

Format

Words: 0

Maks poeng: 10

6(a) **Innleiing**: For å undersøka målenøyaktigheten til ein ny type måleinstrument som skal brukast til å måla farten til bilar blir gjord fleire fartsmålingar på bilar som køyrer i nøyaktig 80 km/t. Vi skal anta at vi gjer n slike målingar og lèt X_i vera differansen mellom målt fart og faktisk fart (altså 80 km/t) i måling nummer $i.$ Dersom målt fart er større enn 80 km/t blir altså X_i positiv, medan X_i blir negativ dersom ein måler farten til å vera mindre enn 80 km/t.

Vi skal føresetja at måleinstrumentet er kalibrert før målingane vi ser på blir gjort slik at $\mathrm{E}[X_i] = 0$ for $i = 1, 2, \ldots, n.$ Dessutan skal vi anta at $X_1, X_2, \ldots, X_n$ er uavhengige og at alle X_i-ene er normalfordelte med same varians $\sigma^2.$

Målenøyaktigheten er dermed gitt av verdien på parameteren $\sigma^2.$

Du kan i denne oppgåva utan bevis nytta at $\mathrm{Var}[X_i^2] = 3\sigma^4.$

Oppgåve: Skriv opp eit uttrykk for rimelighetsfunksjonen for σ^2 i denne situasjonen og bruk dette til å utlede at sannsynlighetsmaksimeringsestimatoren for σ^2 er

$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$

Vis at $\hat{\sigma}^2$ er forventningsrett og at $\mathrm{Var}[\hat{\sigma}^2] = \frac{3\sigma^4}{n}.$

Det blir anbefalt at du løyser denne oppgåva med papir og blyant.

Format

Σ

Words: 0

Maks poeng: 10

6(b) **Oppgåve**: Nytt kjente eigenskapar for normal- og kjikvadratfordelingene, og samanhengar mellom desse fordelingane, til å grunngi at

$$\frac{n\hat{\sigma}^2}{\sigma^2}$$

er kjikvadratfordelt med n fridomsgrader. Angi både kva kjente eigenskapar du nyttar og korleis du nyttar desse i situasjonen vi ser på her.

Utled eit $(1-\alpha)\cdot 100\%$ konfidensintervall for σ^2.

Det blir anbefalt at du løyser denne oppgåva med papir og blyant.

Format

Words: 0

Maks poeng: 10

6(c) **Innleiing**: Konfidensintervallet du utledet i førre punkt er eksakt og gjeld derfor for alle verdiar av n. Vi skal no sjå på korleis sentralgrenseteoremet kan nyttast til å konstruera eit alternativt konfidensintervall for σ^2 som berre er gyldig når tal målingar n er tilstrekkeleg stor.

Oppgåve: Skriv ned antagelsene og resultatet i sentralgrenseteoremet.

Forklar korleis sentralgrenseteoremet kan nyttast til å grunngi at

$$\frac{\hat{\sigma}^2-\sigma^2}{\sigma^2\sqrt{\frac{3}{n}}}$$

er tilnærma standard normalfordelt når n er tilstrekkeleg stor.

Det blir anbefalt at du løyser denne oppgåva med papir og blyant.

Format

Words: 0

Maks poeng: 10

6(d) **Innleiing**: La som vanleg $z_{\frac{\alpha}{2}}$ vera kvantilen slik at $P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$ når Z er standard normalfordelt.

Oppgåve: Nytt resultatet i førre punkt til å visa at

$$\left[\frac{\hat{\sigma}^2}{1+z_{\frac{\alpha}{2}}\sqrt{\frac{3}{n}}}, \frac{\hat{\sigma}^2}{1-z_{\frac{\alpha}{2}}\sqrt{\frac{3}{n}}}\right]$$

er eit tilnærma $(1-\alpha)\cdot 100\%$ konfidensintervall for σ^2 når n er stor.

Det blir anbefalt at du løyser denne oppgåva med papir og blyant.

Format

Words: 0

Maks poeng: 10

6(e) **Innleiing**: Ein har tidlegare nytta ein annan type instrument for å måla farten på bilar. Dette gamle måleinstrumentet har ein mykje erfaring med og ein veit at målenøyaktigheten for dette instrumentet er angitt ved ein verdi σ_0^2.

Vi ønsker no å nytta dei målte fartane med dei nye instrumentet til å avgjera om det er grunnlag for å påstå at målenøyaktigheten er betre med det nye instrumentet enn med det gamle. Dette blir formulert som eit hypotesetestingsproblem med hypotesane

$H_0 : \sigma^2 = \sigma_0^2$ mot $H_1 : \sigma^2 < \sigma_0^2$.

Vi vel å forkasta H_0 dersom $\hat{\sigma}^2 \leq k\sigma_0^2$, der k er ein konstant som må veljast slik at ein får ønskt signifikansnivå.

Oppgåve: Bestem ein verdi for k slik at sannsynet for type I-feil blir (eksakt) lik α. Angi svaret som ein formel for k som funksjon av ein eller fleire av α, n og σ_0^2, og dessutan kvantiler i kjikvadratfordelingen. Hint: Nytt resultatet i (c).

Nytt at $\hat{\sigma}^2$ er tilnærma normalfordelt til å finna ein (tilnærma) sannsyn for type II-feil når $\sigma^2 = \gamma\sigma_0^2$, for ein spesifisert verdi $\gamma < 1$. Angi svaret som ein formel for sannsynet som funksjon av ein eller fleire av av α, n, σ_0^2 og γ, og dessutan kvantiler i kjikvadratfordelingen og den kumulative fordelingsfunksjonen for ein standard normalfordeling.

Det blir anbefalt at du løyser denne oppgåva med papir og blyant.

Format

Words: 0

Maks poeng: 10

7 **Innleiing**: La vera binomisk fordelt med n forsøk og sannsyn for suksess lik p. Ein forventningsrett estimator for p er då

$$\hat{p} = \frac{X}{n}.$$

Anta at vi no er interesserte i $\theta = \mathrm{Var}[X] = np(1-p)$. Som estimator for θ skal vi vurdera

$$\hat{\theta} = n\hat{p}(1-\hat{p}).$$

Oppgåve: Er $\hat{\theta}$ forventningsrett? Grunngi svaret.

Dersom $\hat{\theta}$ ikkje er forventningsrett, foreslå ein modifisert estimator for θ som er forventningsrett.

Det blir anbefalt at du løyser denne oppgåva med papir og blyant.

Format

Words: 0

Maks poeng: 10