i Institutt for matematiske fag

Eksamensoppgåve i TMA4240 Statistikk

Eksamensdato: Torsdag 9. desember

Eksamenstid (frå-til): 09.00-13.00

Hjelpemiddelkode/Tillatne hjelpemiddel: Hjelpemiddelkode C.

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

Fagleg kontakt under eksamen: Håkon Tjelmeland / Sara Martino
Tlf.: 4822 1896 / 9940 3330

ANNAN INFORMASJON:

Skaff deg eit overblikk over oppgåvesettet før du byrjar å svare på oppgåvene.

Generell informasjon:

- I oppgåve 1 skal du berre oppgi korrekt svar (altså utan grunngiving) direkte i Inspera. Følg instruksane i oppgåveteksten om talet på 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 2, 3 og 4 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 deg opp dine eigne meiningar og presiser i svara dine kva for føresetnadar du har lagt til grunn i tolking/avgrensing av oppgåva. Fagleg kontaktperson skal berre kontaktast dersom du meiner det er direkte feil eller manglar i oppgåvesettet. Vend deg til ei eksamensvakt om du ynskjer å kontakte faglærar. Noter gjerne spørsmålet ditt på førehand.

InsperaScan: I oppgåve 2, 3 og 4 skal du levera svaret på ark. Nederst i oppgåva finn du ein sjusifra kode. Fyll inn denne koden øverst til venstre på dei arka du ynskjer å levera. Det er tilrådd å gjere dette undervegs i eksamen. Dersom du behøver tilgang til kodane etter at eksamenstida har gått ut, må du klikke «Vis besvarelse».

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

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 frå/avbroten eksamen: Blir du sjuk under eksamen, eller av andre grunnar ynskjer å levere blankt/avbryte eksamen, gå til “hamburgermenyen” i øvre høgre hjørne og vel «Lever blankt».

Dette kan ikkje angrast sjølv om prøven framleis er open.

Tilgang til svara dine: Etter eksamen finn du svara dine i arkivet i Inspera. Merk at det kan ta ein virkedag før eventuelle handteikningar vert tilgjengelege i arkivet.

1(a) **Innleiing**: La X og Y vere to uavhengige og normalfordelte stokastiske variablar, der X har forventningsverdi lik 2 og standardavvik lik 1, og Y har forventningsverdi lik -1 og standardavvik lik 3.

Oppgåve: Rekn ut følgjande storleikar. Angi svara med tre desimalar etter komma.

- $P(X \geq 3) =$
- $P(2X - Y \leq 0) =$
- $P\left(X^2 \leq \frac{1}{4}\right) =$

Maks poeng: 5

1(b) **Innleiing**: Anta at det er 55 studentar som tar eit emne ved NTNU. Av desse 55 studentane er det 25 gutar og 30 jenter. I dette emnet skal det no bli oppnemnt ei referansegruppe beståande av 5 studentar.

Oppgåve: Finn følgjande tal. Oppgi svara som heiltal.

- Kor mange ulike referansegrupper kan ein danna i dette emnet?
- Kor mange ulike referansegrupper kan ein danna i dette emnet dersom ein krev at gruppa skal ha minst to medlemmer av kvart kjønn?

Maks poeng: 5

1(c) **Innleiing**: La X vera binomisk fordelt med $n = 25$ forsøk og sannsyn for suksess $p = 0.33.$ Som kjent kan ein binomisk fordeling med god approksimasjon tilnærmast med ei normalfordeling når $np \geq 5$ og $n(1-p) \geq 5,$ eit vilkår som er oppfylt for fordeling til $X.$

Oppgåve: Nytt normalapproksimasjon til å finna (tilnærma) verdiar for følgjande sannsyn. Husk å nytta kontinuitetskorreksjon når du reknar ut dei tilnærma verdiane. Angi svara med tre desimalar etter komma.

- $P(X < 5) =$ ☐
- $P(X > 10) =$ ☐
- $P(X > 10 | X \geq 4) =$ ☐

Maks poeng: 5

1(d)

Innleiing: La X og Y vera to diskrete stokastiske variablar med simultanfordeling $P(X = x, Y = y)$ som gitt i tabellen under:

$x \setminus y$	0	1	2
0	0.22	0	0.31
1	0.32	0.15	0

Oppgåve: Rekn ut følgjande storleikar. Angi svara med tre desimalar etter komma.

- $P(X = 1) =$ ☐
- $P(Y = 0 | X = 0)$ = ☐
- $E(2XY) =$ ☐

Maks poeng: 5

2 **Innleiing**: La X vera ein kontinuerleg fordelt stokastisk variabel. Anta vidare at X har fordeling

$$f(x) = \begin{cases} \frac{2x}{10} e^{-x^2/10} & \text{for } x \geq 0, \\ 0 & \text{elles.} \end{cases}$$

Oppgåve: Finn følgjande storleikar:

- $P(X \geq 2)$
- $P(X^2 - 4X \leq -1)$
- Medianen til X.

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

Format

Words: 0

Maks poeng: 10

3(a) **Innleiing**: Levetida til ein elektronisk komponent T blir antatt å vera eksponensialfordelt med parameter λ. Sannsynstettleiken til T er dermed gitt som

$$f(t;\lambda) = \begin{cases} \frac{1}{\lambda}\exp(-\frac{t}{\lambda}) & \textbf{for } t \geq 0, \\ 0 & \textbf{elles.} \end{cases}$$

La $T_1, T_2, \dots, T_n$ vera eit tilfeldig utval frå denne fordelinga.

Oppgåve:

- Vis først at $\frac{2T_i}{\lambda}$ er gammafordelt.
- Bruk så dette til å visa at $V = \frac{2\sum_{i=1}^{n} T_i}{\lambda}$ er χ^2-fordelt med $2n$ fridomsgrader.

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

Format B I U x_2 x^2 I_x Ω Σ

Words: 0

Maks poeng: 10

3(b) **Innleiing**: Anta at 20 slike elektroniske komponentar er vorte testa og gjennomsnittet av dei observerte levetidene er $\bar{t} = 10.5$.

Oppgåve: Utlei eit 90% konfidensintervall for λ og bruk observert verdi for $\bar{t}$ til å rekna ut intervallet numerisk.

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

Format

Words: 0

Maks poeng: 10

4(a) **Innleiing**: Anta at du er vorte bedt om å undersøka talet på ulykker på byggeplassar i Noreg og har samla inn informasjon om ulykker frå $n = 34$ byggeplassar. Anta at ulykker på ulike byggeplassar skjer uavhengig av kvarandre og at ulykker på byggeplass nummer i skjer ifølgje ein poissonprosess med intensitet λm_i, der m_i er talet på tilsette på denne byggeplassen og λ er ein parameter som blir antatt felles for alle byggeplassar. La X_i vere talet på ulykker på byggeplass nummer i i løpet av $t = 1$ år. Dermed er X_i poissonfordelt med forventningsverdi lik λm_i og har punktsannsyn

$$f(x_i) = \frac{(\lambda m_i)^{x_i}}{x_i!} e^{-\lambda m_i}, x_i = 0, 1, 2, \ldots$$

Oppgåve: For $\lambda = 0.020$ og $m_i = 50$, svar på følgjande spørsmål.

- Kva er sannsynet for at det blir minst ei ulykke på byggeplass nummer i?
- Gitt at det er minst ei ulykke på byggeplass nummer i, kva er sannsynet for at det er meir enn ei ulykke på denne byggeplassen?
- Gitt at det er minst ei ulykke på byggeplass nummer i, kva er sannsynet for at det er maksimalt tre ulykker på denne byggeplassen?

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

Format

Words: 0

Maks poeng: 10

4(b) **Innleiing:** I resten av oppgåva skal vi anta at verdien til parameteren λ er ukjent og at vi ønsker å estimera denne frå observerte verdiar $x_1, x_2, \ldots, x_n$ av talet på ulykker på kvar byggeplass. Vi skal anta at talet på tilsette på kvar byggeplass, $m_1, m_2, \ldots, m_n$, er kjent.

Oppgåve: Finn rimelighetsfunksjonen $L(\lambda)$ og nytt denne til utleie sannsynsmaksimeringsestimatoren, $\widehat{\lambda}$, for λ. Vis spesielt at denne estimatoren kan skrivast som

$$\widehat{\lambda} = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n m_i}.$$

Vis vidare at $\widehat{\lambda}$ er ein forventningsrett estimator for λ og at

$$\mathrm{Var}[\widehat{\lambda}] = \frac{\lambda}{\sum_{i=1}^n m_i}.$$

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

Format

Words: 0

Maks poeng: 10

4(c) **Innleiing**: For 15 år sidan vart det gjennomført ei grundig undersøking om talet på ulykker på norske byggeplassar, og ein fann då at verdien på parameteren λ var $0.020.$ Du ønsker no å nytta observerte verdiar for X_i-ane til å vurdera om det er grunnlag for å påstå at talet på ulykker på norske byggeplassar er lågare i dag enn det var for 15 år sidan.

Du kan i resten av denne oppgåva føresetja (utan bevis) at

$$\frac{\widehat{\lambda}-\mathrm{E}[\widehat{\lambda}]}{\sqrt{\mathrm{Var}[\widehat{\lambda}]}}$$

er tilnærma standard normalfordelt.

Oppgåve: Formuler hypotesane H_0 og H_1 for denne situasjonen, vel ein testobservator, angi ein avgjerdsregel og bestem testens p-verdi (tilnærma) når dei observerte verdiane er gitt i tabellen under.

Basert på denne p-verdien, vil du seia at det er grunnlag for å påstå at talet på ulykker i dag er lågare enn han var for 15 år sidan?

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
x_i	1	0	2	3	0	0	1	2	0	0	2	0	1	0	0	1	1
m_i	28.8	55.2	49.4	55.9	34.8	34.6	28.4	77.0	81.1	29.3	62.8	22.6	106.3	44.9	27.8	49.8	73.1
i	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
x_i	0	0	1	2	2	1	1	0	2	3	0	2	1	0	1	1	2
m_i	43.8	21.7	64.5	77.6	79.6	18.2	69.5	23.1	63.8	56.2	65.0	28.9	52.8	53.4	49.0	40.6	115.0

Det blir oppgitt at desse verdiane gir $\sum_{i=1}^{n} x_i = 33$ og $\sum_{i=1}^{n} m_i = 1784.4.$

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

Format B I U x_2 x^2 I_x Ω Σ

Words: 0

Maks poeng: 10

4(d) **Innleiing**: Anta no at ein har moglegheit for å innhenta informasjon om talet på ulykker på dei $n = 34$ byggeplassane over ein lengre periode enn $t = 1$ år, og at du ønsker å nytta dette til å redusera sannsynet for type-II feil i hypotesetesten du formulerte i (d).

Du kan føresetja at ein i hypotesetesten bruker signifikansnivå $\alpha = 0.05,$ og at verdiane til $m_1, m_2, \ldots, m_n$ er konstant lik verdiane angitt i tabellen i oppgåve (c).

Oppgåve: Dersom den sanne verdien for λ i dag er $0.018,$ for kor lang tid t må du innhenta informasjon om talet på ulykker for at sannsynet for å oppdaga at H_0 er feil skal vera minst $0.2.$(Hint: Finn først uttrykk for sannsynsmaksimeringsestimatoren for λ når X_i er talet på ulykker ved byggeplass nummer i i løpet av t år.)

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

Format

Words: 0

Maks poeng: 10

4(e) **Innleiing**: Som eit alternativ til å nytta sannsynsmaksimeringprinsippet for å konstruera ein estimator for λ er det foreslått å nytta ein estimator på forma

$$\tilde{\lambda} = \sum_{i=1}^{n} b_i X_i,$$

der $b_1, b_2, \ldots, b_n$ er konstantar (som kan vera funksjonar av $m_1, m_2, \ldots, m_n$).

Oppgåve: Vis at estimatoren $\tilde{\lambda}$ blir best ved å la alle b_i-ane vera like og gitt som

$$b_i = \frac{1}{\sum_{j=1}^{n} m_j}.$$

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

Format

Words: 0

Maks poeng: 10