Institutt for matematiske fag

Eksamensoppgåve i TMA4240 Statistikk

Eksamensdato: Torsdag 9. desember

Eksamenstid (frå-til): 15.00-19.00

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 vera uavhengige og normalfordelte stokastiske variablar. La X ha forventningsverdi lik -2 og standardavvik lik $2,$ og la Y ha forventningsverdi lik 1 og standardavvik lik $1.$

Oppgåve: Finn følgjande sannsyn. Oppgi svara med tre desimalar etter komma.

- $P(X > 0) =$ ☐
- $P(2X - Y \geq -1) =$ ☐
- $P(Y \geq X) =$ ☐

Maks poeng: 5

1(b) **Innleiing**: Anta at vi har ei urne med 15 raude og 20 blå kuler. Frå denne urna trekkar vi utan tilbakeleggjing ut 6 kular, og antar at trekningane blir gjorde slik at alle kular har like stort sannsyn for å bli trekt ut. Vi er så interesserte i kva 6 kular vi har trekt ut.

Oppgåve: Finn følgjande tal. Oppgi svara som heiltal.

- Når vi reknar kvar kule som unik, på kor mange måtar kan ein trekka ut dei 6 kulane? ☐
- Når vi reknar kvar kule som unik, på kor mange måtar kan ein trekka ut dei 6 kulane slik at det er trekt ut minst 2 kular av kvar farge? ☐

Maks poeng: 5

1(c) **Innleiing**: La $X_1, X_2, \ldots, X_{50}$ vera eit tilfeldig utval frå ei sannsynsfordeling med forventningsverdi $\mu = \mathrm{E}[X_i] = 3.5$ og varians $\sigma^2 = \mathrm{Var}[X_i] = 1.5$. La $\bar{X}$ vere gjennomsnittet av X_i-ane, dvs

$\bar{X} = \frac{1}{50}\sum_{i=1}^{50} X_i.$

Oppgåve: Bruk sentralgrenseteoremet til å finna tilnærma verdiar for følgjande sannsyn. Oppgi svara med tre desimalar etter komma.

- $P(\bar{X} \leq 4.0) =$ ☐
- $P\left(\sum_{i=1}^{50} X_i \leq 185\right) =$ ☐
- $P\left(\bar{X} \leq 4.0 \middle| \sum_{i=1}^{10} X_i = 48.2\right) =$ ☐

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
0	0.4	0.1
1	0.3	0
2	0	0.2

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

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

Maks poeng: 5

2 **Innleiing**: La X vera ein kontinuerleg fordelt stokastisk variabel som er gammafordelt med parametrar $\alpha = 2$ og $\beta = \frac{1}{2}$. Sannsynstettleiken til X er dermed gitt som

$$f(x) = \begin{cases} 4xe^{-2x} & \text{for } x \geq 0, \\ 0 & \text{elles.} \end{cases}$$

Oppgåve: Finn følgjande storleikar.

- $P(X \leq 2)$
- $P(X^2 - 4X < -3)$
- $\mathrm{E}[\sqrt{X}]$

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

Format

Words: 0

Maks poeng: 10

3(a) **Innleiing**: Det har brote ut eit virus i samfunnet. Myndigheitene ønsker å finna kva for ein del av befolkninga som er smitta ved å testa eit tilfeldig utval av N personar.

La p vera sannsynet for at ein tilfeldig utvald person frå befolkninga er smitta. Anta at testen ikkje er heilt påliteleg, sannsynet for eit "falskt negativt" resultat er 0.02. Dette betyr at sannsynet for at ein sjuk person testar positiv er 0.98. Du kan anta at sannsynet for eit "falskt positivt" resultat, det vil seia at ein frisk person testar positivt, er null.

La X vera ein stokastisk variabel som er lik talet på dei N personane i utvalet som testar positivt.

Oppgåve:

- Forklar kvifor X er binomisk fordelt med parametrar N og $0.98p$.
- Vis at $\hat{p} = \frac{X}{0.98N}$ er ein forventningsrett estimator for p og rekn ut standardavviket til denne estimatoren.

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

Format

Words: 0

Maks poeng: 10

3(b) **Innleiing:** Myndigheitene tester $N = 1000$ personar og for $x = 55$ av desse er testen positiv.

Oppgåve: Utlei eit 95% konfidensinterval for p basert på desse observasjonane.

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

Format

Words: 0

Maks poeng: 10

4(a) **Innleiing**: Levetida (målt månader) til nokon typar mekaniske komponentar har vist seg å følgja ei fordeling med sannsynstettleik

$$f(x) = \begin{cases} \frac{1}{2\theta}\frac{1}{\sqrt{x}}\exp\left\{-\frac{\sqrt{x}}{\theta}\right\} & \text{for } x \geq 0, \\ 0 & \text{elles,} \end{cases}$$

der $\theta > 0$ er ein parameter som beskriv kvaliteten av denne typen komponentar. Tilhøyrande kumulativ fordelingsfunksjon er

$$F(x) = \begin{cases} 1 - \exp\left\{-\frac{\sqrt{x}}{\theta}\right\} & \text{for } x \geq 0, \\ 0 & \text{elles.} \end{cases}$$

I heile denne oppgåva skal vi dessutan anta at levetida til ulike komponentar er uavhengige av kvarandre.

[Hint: Du kan i nokon av punkta (a)-(e) i denne oppgåva få bruk for å nytta at

$$\int_0^\infty x^\alpha \exp\left\{-\frac{\sqrt{x}}{\theta}\right\}dx = 2\theta^{2\alpha+2}\Gamma(2\alpha+2) \text{ for } \alpha > -1, \theta > 0,$$

der $\Gamma(\alpha)$ er gammafunksjonen.]

Oppgåve: For $\theta = 2.5,$ svar på følgjande spørsmål.

- Kva er sannsynet for at ein slik komponent framleis vil funksjonera etter 15 månader?
- Viss ein veit at ein slik komponent framleis funksjonerer etter 15 månader, kva er sannsynet for at denne komponenten framleis vil funksjonera etter 20 månader?
- Kva er forventa levetid for ein slik komponent?

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

Format

Words: 0

Maks poeng: 10

4(b) **Innleiing**: For å undersøka kvaliteten på ein ny type av slike komponentar har ein prøvd ut $n = 25$ av desse. La $X_1, X_2, \ldots, X_n$ vera levetidene, slik at dette altså er eit tilfeldig utval frå fordelinga spesifisert i (a). Dei observerte levetidene, $x_1, x_2, \ldots, x_n$, er:

1.01, 7.10, 9.41, 0.79, 40.06, 1.90, 2.77, 26.91, 4.04, 2.79, 6.58, 41.46, 5.68, 2.37, 5.08, 1.59, 1.00, 0.07, 1.44, 0.27, 118.22, 0.07, 5.21, 2.85, 101.87.

Det blir oppgitt at dette gir $\sum_{i=1}^{n} \sqrt{x_i} = 70.493.$

Oppgåve: Vis at sannsynsmaksimeringsestimatoren for θ er

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} \sqrt{X_i}.$$

Er $\hat{\theta}$ en forventningsrett estimator for θ? Grunngi svaret.

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

Format

Words: 0

Maks poeng: 10

4(c) **Innleiing**: For $i = 1, 2, \ldots, n$, definer

$$Z_i = \frac{2}{\theta}\sqrt{X_i}.$$

Oppgåve: Vis ved hjelp av transformasjonsformelen at Z_i er χ^2-fordelt med 2 fridomsgrader.

Bruk så dette til å grunngi at

$$\frac{2n\hat{\theta}}{\theta}$$

er χ^2-fordelt med $2n$ fridomsgrader.

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

Format

Words: 0

Maks poeng: 10

4(d) **Innleiing**: Ein vurderer å ta den nye typen av komponentar i bruk i staden for ein gammal type som har vore lenge i bruk. For den gamle typen komponentar som har vore lenge i bruk veit ein at verdien til θ er 2.6. Ein ønsker derfor å undersøka om dei observerte levetidene gitt i (b), som er for komponentar av den nye typen, gir grunnlag for å påstå at forventa levetid for den nye typen komponentar er større enn forventa levetid for den gamle typen komponentar.

For å svara på oppgåva kan du nytta tabellen på nest siste side av oppgåvesettet. Tabellen inneheld sannsyn i ein χ^2-fordeling med 50 fridomsgrader.

Oppgåve: Formuler situasjonen skildra over som eit hypotesetestingsproblem. Formuler hypotesane H_0 og H_1, angi ein testobservator og ein tilhøyrande avgjerdsregel, og bestem p-verdien så godt du kan ut frå tabellen på nest siste side i oppgåvesettet.

Basert på denne p-verdien, vil du seia at det er grunnlag for å påstå at forventa levetid for den nye typen komponentar er større enn forventa levetid for den gamle typen komponentar? Grunngi svaret.

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

Format

Σ

Words: 0

Maks poeng: 10

4(e) **Innleiing**: Anta at du får høve til å prøva ut fleire komponentar av den nye typen, og at du ønsker å nytta dette til å redusera sannsynet for type II-feil i testen formulert i (d).

Du kan føresetja at ein i hypotesetesten bruker signifikansnivå $\alpha = 0.05.$

For å svara på oppgåva kan du nytta tabellen på den siste sida av oppgåvesettet. Tabellen inneheld kvantiler i ein χ^2-fordeling for ulike fridomsgrader.

Oppgåve: Dersom den sanne verdien for θ for den nye typen komponentar er 3.4, kor mange levetider for komponentar av den nye typen må ein observera for at sannsynet for å oppdaga at H_0 er feil skal vera minst 0.50?

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

Format | B I U | Σ

Words: 0

Maks poeng: 10

Sannsyn i χ^2-fordeling med 50 fridomsgrader

$P(X^2 \leq z)$ når $X^2 \sim \chi^2_{50}$

z	$P(X^2 \leq z)$
45.0	0.3262
45.5	0.3457
46.0	0.3654
46.5	0.3854
47.0	0.4055
47.5	0.4257
48.0	0.4460
48.5	0.4663
49.0	0.4865
49.5	0.5066
50.0	0.5266
50.5	0.5464
51.0	0.5659
51.5	0.5852
52.0	0.6041
52.5	0.6226
53.0	0.6408
53.5	0.6585
54.0	0.6758
54.5	0.6927
55.0	0.7090
55.5	0.7248
56.0	0.7401
56.5	0.7548
57.0	0.7691
57.5	0.7827
58.0	0.7958
58.5	0.8084
59.0	0.8204
59.5	0.8318
60.0	0.8428

i

Kvantiler i χ^2-fordelingen

$$P(X^2 > \chi^2_{\alpha,\nu}) = \alpha$$

$\nu \backslash \alpha$	**0.95**	**0.50**	**0.05**
60	43.188	59.335	79.082
61	44.038	60.335	80.232
62	44.889	61.335	81.381
63	45.741	62.335	82.529
64	46.595	63.335	83.675
65	47.450	64.335	84.821
66	48.305	65.335	85.965
67	49.162	66.335	87.108
68	50.020	67.335	88.250
69	50.879	68.334	89.391
70	51.739	69.334	90.531
71	52.600	70.334	91.670
72	53.462	71.334	92.808
73	54.325	72.334	93.945
74	55.189	73.334	95.081
75	56.054	74.334	96.217
76	56.920	75.334	97.351
77	57.786	76.334	98.484
78	58.654	77.334	99.617
79	59.522	78.334	100.749
80	60.391	79.334	101.879
81	61.261	80.334	103.010
82	62.132	81.334	104.139
83	63.004	82.334	105.267
84	63.876	83.334	106.395
85	64.749	84.334	107.522
86	65.623	85.334	108.648
87	66.498	86.334	109.773
88	67.373	87.334	110.898
89	68.249	88.334	112.022