i **Institutt for matematiske fag, NTNU**

Eksamensoppgåve i **TMA4240 Statistikk**

Fagleg kontakt under eksamen: Sara Martino og Håkon Tjelmeland
Tlf: 44903330, 48221896

Eksamensdato: 30. november 2019
Eksamenstid (frå-til): 09:00 - 13:00
Hjelpemiddelkode/Tillatne hjelpemiddel: Hjelpemiddelkode C.

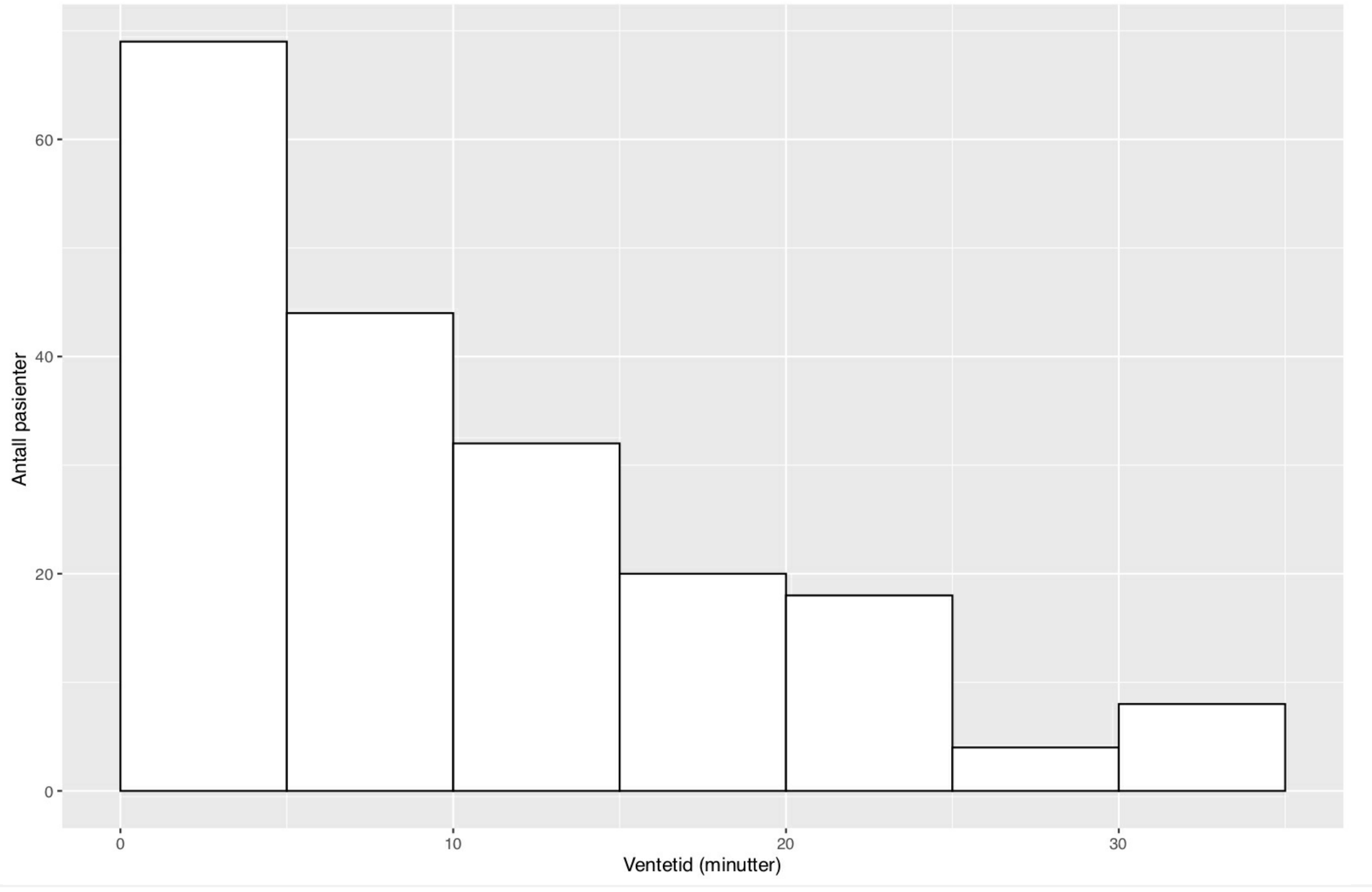
- *Tabeller og formler i statistikk* (Fagbokforlaget)
- Eit gult ark (A5 med stempel) med eigne handskrivne formlar og notat
- Bestemd, enkel kalkulator

Annan informasjon:

- For oppgåvene som ikkje er fleirvalsoppgåver, må alle svar grunngivast og alle svar må innehalda mellomrekning slik at det er heilt klart korleis ein har tenkt.
- Vekt ved sensur for kvar deloppgåve er angitt i oppgåvesettet.

Målform/språk: Nynorsk

1(a)



Fleirvalsoppgåver

Innleiing: Histogrammet over viser frekvensfordelinga for ventetider for pasientar på eit tannlegekontor.

Oppgåve: Kva er (omtrentleg) medianen av desse ventetidene?

Vel eitt alternativ

- 2.50
- 7.25
- 12.25
- 15.00
- 17.50

Vekt ved sensur: 4 %.

Maks poeng: 4

1(b) **Innleiing**: Du skal utføra ein einsidig hypotesetest der nullhypotesen er at forventningsverdien er lik 532 og alternativ hypotese er at forventningsverdien er mindre enn 532. Anta at gjennomsnittet av observasjonane dine er 529 og at p-verdien er 0.01.

Oppgåve: Kva for av følgjande utsegner er då korrekte?

Vel eitt eller fleire alternativ

- ☐ Sannsynet for at forventningsverdien er mindre enn 529 er 0.01.
- ☐ Viss forventningsverdien er 532 er sannsynet for å observera eit gjennomsnitt som er mindre enn eller lik 529 lik 0.01.
- ☐ Sannsynet for at forventningsverdien er mindre enn eller lik 532 er 0.01.
- ☐ Viss signifikansnivået er 0.05, vil ein ikkje forkasta H_0.
- ☐ Ingen av fråsegnene over er korrekte.

Vekt ved sensur: 4 %.

Maks poeng: 4

1(c) **Innleiing**: Ein produsent av ballongar hevdar at delen, p, av ballongane som sprekk når dei blir blåsne opp til å ha ein diameter på $30\ cm$, er mindre enn 0.05. Ein del kundar har klaga på at delen av ballongane som sprekk er høgare enn han hevdar.

Oppgåve: Dersom desse kundane ønsker å gjennomføra ei undersøking for å testa påstanden til produsenten, kva for av følgjande hypotesar vil vera riktige?

Vel eitt eller fleire alternativ

- ☐ $H_0 : p \neq 0.05$ mot $H_1 : p = 0.05$.
- ☐ $H_0 : p = 0.05$ mot $H_1 : p \neq 0.05$.
- ☐ $H_0 : p = 0.05$ mot $H_1 : p < 0.05$.
- ☐ $H_0 : p = 0.05$ mot $H_1 : p > 0.05$.
- ☐ $H_0 : p < 0.05$ mot $H_1 : p = 0.05$.

Vekt ved sensur: 4 %.

Maks poeng: 4

1(d) **Innleiing**: Anta at ein i ei undersøking fann eit 95 %-konfidensintervall for delen av nordmenn som trenar regelmessig til å gå frå 0.29 til 0.37.

Oppgåve: Kva for av følgjande fråsegn er då FEIL?

Vel eitt eller fleire alternativ

- ☐ Det er rimeleg å seia at meir enn 25 % av den norske befolkninga trenar regelmessig.
- ☐ Det er rimeleg å seia at meir enn 40 % av den norske befolkninga trenar regelmessig.
- ☐ Ein hypotese om at meir enn 33 % av den norske befolkninga trenar regelmessig kan ikkje forkastast.
- ☐ Det er rimeleg å seia at mindre enn 40 % av den norske befolkninga trenar regelmessig.

Vekt ved sensur: 4 %.

Maks poeng: 4

1(e) **Innleiing**: La X og Y vera to uavhengige stokastiske variablar som begge er geometrisk fordelte med parameter p, dvs

$$X \sim g(x;p) = p(1-p)^{x-1}; x = 1,2,\ldots \text{ og } Y \sim g(y;p) = p(1-p)^{y-1}; y = 1,2,\ldots$$

La $Z = \min(X,Y)$. Du kan i denne oppgåva få bruk for formelen for ei geometrisk rekke,

$$\sum_{k=0}^{n} a^k = \frac{1-a^{n+1}}{1-a}$$

Oppgåve: Kva er punktsannsynet $f_Z(z)$ for Z?

Vel eitt alternativ

- ○ $f_Z(z) = (1-p)^{2(z-1)}p(2-p); z = 1,2,\ldots$
- ○ $f_Z(z) = 1-(1-p)^{2z}; z = 1,2,\ldots$
- ○ $f_Z(z) = (1-(1-p)^z)^2; z = 1,2,\ldots$
- ○ $f_Z(z) = p^2(1-p)^{2(z-1)}; z = 1,2,\ldots$
- ○ $f_Z(z) = -2(1-p)^{2z}\ln(1-p); z = 1,2,\ldots$

Vekt ved sensur: 4 %.

Maks poeng: 4

2 Normalfordeling

Innleiing: La X og Y vera to uavhengige og normalfordelte stokastiske variablar. Anta at X har forventningsverdi lik 0 og standardavvik lik 2, medan Y har forventningsverdi lik 1 og standardavvik lik 1.

Oppgåve:

- Skisser sannsynstettleikane til X og Y i eit felles plott.
- Finn følgjande sannsyn,

$$P(X \leq 1), \quad P(Y \geq -1) \quad \text{og} \quad P(X - Y \leq 0)$$

Skriv svaret ditt her...

Format B I U x₂ x² Ω Σ

Words: 0

Vekt ved sensur: 10 %.

Maks poeng: 10

3(a) Sigdcelleanemi

Innleiing: Sigdcelleanemi er ein alvorleg sjukdom som forårsakar anemi, dvs låg blodprosent. Sigdcelleanemi er arveleg, og eit barn får sjukdommen dersom det arvar eit bestemt recessivt gen (a) frå både mor og far. Barnet får ikkje sjukdommen dersom det arvar det dominante genet (A) frå minst ein av foreldra sine.

Ein person har altså enten genotype AA, Aa eller aa. Personar med genotype aa har sigdcelleanemi, medan personar med AA eller Aa har ikkje sjukdommen. Eit barn arvar eit gen frå kvar av foreldra sine. Viss ein forelder har genotype Aa, vil eit barn arva enten a eller A frå denne forelderen med sannsyn 0.5 for kvar. Eit barn arvar gen frå mor og far uavhengig av kvarandre.

Ein seier at personar som har genotype Aa er berarar av sjukdommen, dei har ikkje sjølv sigdcelleanemi men kan få barn som har sjukdommen.

Vurder no eit par der verken mannen eller kvinna har sigdcelleanemi, men vi veit ikkje om dei er berarar av sjukdommen. Anta at mannen og kvinna kvar har eit sannsyn på $8\ \%$ for å vera berarar av sjukdommen (dette er delen av den afroamerikanske befolkninga som er berarar) og at mannen og kvinna er berarar eller ikkje uavhengig av kvarandre.

Vi skal i denne oppgåva rekna på sannsynet for at dette paret får barn som har sigdcelleanemi eller som er berarar av sjukdommen. Definer hendingane

- M: Mannen er bærar av sigdcelleanemi,
- K: Kvinna er bærar av sigdcelleanemi,
- D: Det første barnet paret får, har sigdcelleanemi, og
- B: Det første barnet paret får, er bærar av sigdcelleanemi.

Oppgåve:

- Teikn opp hendingane M, K, D og B i eit venndiagram.
- Rekn ut sannsynet for at det første barnet paret får vil ha sigdcelleanemi.
- Rekn ut sannsynet for at det første barnet paret får vil vera berar av sigdcelleanemi.

Skriv svaret ditt her...

Format

Words: 0

Vekt ved sensur: 10 %.

Maks poeng: 10

3(b) **Innleiing**: Vi skal så anta at paret allereie har fått eit barn som ikkje har sigdcelleanemi, og at dei no planlegg å få eit barn til.

Oppgåve:

- Kva er då sannsynet for at det andre barnet vil ha sigdcelleanemi?
- Kva er då sannsynet for at det andre barnet vil vera berar av sigdcelleanemi?

Skriv svaret ditt her...

Format

Words: 0

Vekt ved sensur: 10 %.

Maks poeng: 10

4(a) **Binomisk fordelte variablar**

Innleiing: Anta at vi har to uavhengige stokastiske variablar X og Y, der $X \sim b(x; n, p)$ og $Y \sim b(y; n, 2p)$. X er altså talet på suksessar i n uavhengige forsøk der kvart forsøk har sannsyn p for å gi suksess, medan Y er talet på suksessar i n andre uavhengige forsøk der kvart forsøk gir suksess med sannsyn $2p$.

Oppgåve:

- Når $n = 12$ og $p = 0.2$, finn følgjande sannsyn,

$$P(X \leq 3), \;\; P(Y \geq 4 | X \leq 3) \;\; \text{og} \;\; P(X + Y \leq 1).$$

Skriv svaret ditt her...

Format B I U x_2 x^2 I_x Ω Σ

Words: 0

Vekt ved sensur: 10 %.

Maks poeng: 10

4(b) **Innleiing**: I resten av oppgåva skal vi anta at verdien av parameteren $p \in [0, 0.5]$ er ukjend og skal estimerast basert på X og Y. Følgjande tre estimatorar er foreslått,

$$\hat{p} = \frac{X+Y}{2n}, \ \tilde{p} = \frac{X+Y}{3n} \ \text{ og } \ p^{\star} = \frac{X}{2n} + \frac{Y}{4n} \tag{1}$$

Oppgåve:

- Kva for ein av dei tre estimatorane vil du føretrekka? Grunngi svaret.

Skriv svaret ditt her...

Format B I U x_2 x^2 I_x Ω Σ

Words: 0

Vekt ved sensur: 10 %.

Maks poeng: 10

4(c) **Innleiing**: Anta at vi òg ønsker å estimera p ved å nytta sannsynsmaksimeringsprinsippet, framleis basert på X og Y.

Oppgåve:

- Finn eit uttrykk for log-rimelegheitsfunksjonen for p, $l(p)$.
- Avgjer kva estimatet for p blir ved å nytta sannsynsmaksimeringsprinsippet når $n = 25$, og ein observerer $x = 3$ og $y = 8$.

Skriv svaret ditt her...

Format B I U x_2 x^2 I_x Ω Σ

Words: 0

Vekt ved sensur: 10 %.

Maks poeng: 10

4(d) **Innleiing**: Anta så at vi ønsker å nytta X og Y til å teste

$$H_0 : p = 0.2 \text{ mot } H_1 : p > 0.2.$$

Vi skal i resten av oppgåva anta at n er stor nok til at ein med god approksimasjon kan anta at X og Y er normalfordelte.

Oppgåve:

- Ta utgangspunkt i den av dei tre estimatorane i (1) du fann var best, vel en testobservator og finn eit (tilnærma) forkastingskriterium når signifikansnivået blir valt som $\alpha = 0.05$. Viss du ikkje konkluderte kva for ein av dei tre estimatorane i (1) som var best, kan du sjølv velja kva for ein av estimatorane du vil ta utgangspunkt i.

Skriv svaret ditt her...

Format B I U x₂ x² Ω Σ

Words: 0

Vekt ved sensur: 10 %.

Maks poeng: 10

4(e) **Oppgåve**:

- Viss ein bruker signifikansnivå $\alpha = 0.05$ i hypotesetesten over, kor stor må n veljast for at ein skal ha minst sannsyn 0.9 for å oppdaga at H_0 er feil når $p = 0.25$?

Skriv svaret ditt her...

Format

Words: 0

Vekt ved sensur: 10 %.

Maks poeng: 10