

Institutt for matematiske fag

Eksamensoppgave i **TMA4240/TMA4245 STATISTIKK**

Faglig kontakt under eksamen: Gunnar Taraldsen

Tlf: 46432506

Eksamensdato: august 2018

Eksamenstid (fra–til): 09:00 – 13:00

Hjelpemiddelkode/Tillatte hjelpemidler: Hjelpemiddelkode C:

- Tabeller og formler i statistikk, Akademika
- Ett gult ark (A4 med stempel) med egne håndskrevne formler og notater
- Bestemt, enkel kalkulator

Annen informasjon:

Alle svar må begrunnes.

Du må ha med nok mellomregninger til at tenkemåten din klart fremgår.

Oppgaven består av 25 delpunkter som har lik vekt ved sensur.

Målform/språk: bokmål

Antall sider: 4

Antall sider vedlegg: 0

Kontrollert av:

Informasjon om trykking av eksamensoppgave

Originalen er:

1-sidig ☐ **2-sidig** ☒

sort/hvit ☒ **farger** ☐

skal ha flervalgskjema ☐

Dato

Sign

Oppgave 1 Karoline er en flittig bruker av sosiale medier og ønsker å treffe nye personer. Anta at Karoline hver dag snakker med en ny person på sosiale medier og foreslår at de skal møtes den samme dagen. Dessverre for Karoline er det ikke alle som vil møtes. La X være antall personer Karoline må spørre før det første møtet og la p være sannsynligheten for at en tilfeldig valgt person vil møte Karoline. Anta i de neste tre punkter at $p = 0.3$.

- a) Gi en begrunnelse for at X er geometrisk fordelt med sannsynlighetstetthet

$$f(x | p) = P(X = x) = p(1 - p)^{x-1} \quad (1)$$

- b) Hva er sannsynligheten for at det går tre dager eller flere før det første møtet?

- c) Beregn $P(X \leq 4 | X > 2)$.

- d) Gi en tolkning av $P(X \leq 4 | X > 2)$.

- e) Anta at Karoline i løpet av 14 dager observerer x verdiene 4, 2, 3, 5. Utled og beregn rimelighetsestimater* $\hat{p}$.

- f) Forklar ved et annet argument hvordan en binomisk modell kan brukes til å begrunne estimater $\hat{p}$.

*sannsynlighetsmaksimeringsestimatet

Oppgave 2 En beslutning om å gi en fartsbot, som potensielt kan gi grunnlag for en straffesak, må kun tas om det er overveiende sannsynlig at fartsgrensen faktisk er overskredet. En type Doppler radar brukt av politiet til å måle farten til kjøretøy gir en fart V som er normalfordelt med forventningsverdi w og standardavvik γw når kjøretøyet har reell fart w . Den relative feilen til radaren er $\gamma = 2\%$.

- a) Anta at politiet måler en fart $v = 135$ km/t. Estimer den reelle farten w .
- b) Finn et estimat u til standardavviket til estimatoren brukt i forrige punkt.
- c) I følge det internasjonale standardiseringsorganet ISO skal alle målinger rapporteres med en tilhørende usikkerhet. Den beregnede u verdien i forrige punkt betegnes av ISO som standard usikkerhet til målingen. Gi en kort begrunnelse for hvorfor det er rimelig å kalle dette en standard usikkerhet.
- d) Anta at fartsgrensen er 90 km/t. Kjøretøyfart w på 131 km/t eller høyere skal gi tapt førerkort for føreren. Skal en fører med målt fart $v = 135$ km/t tape førerkortet?

Svaret skal begrunnes ved å beregne en p-verdi tilsvarende nullhypotesen $H_0 : w < 131$ km/t og den alternative hypotesen $H_1 : w \geq 131$ km/t. Bruk signifikansnivå $\alpha = 0.1\%$.
- e) En test er gitt ved å forkaste H_0 dersom $v > v_m$. Bestem v_m slik at testens signifikansnivå blir $\alpha = 0.1\%$.

Styrkefunksjonen h til testen er definert ved $h(w) = P(V > v_m)$, d.v.s. lik sannsynligheten for å forkaste H_0 som funksjon av modellparameteren w . Her avhenger styrken h av parameteren w fordi fordelingen til V avhenger av w .

- f) Skisser grafen til styrkefunksjonen til testen.
Tips: Beregn for verdiene $w = 137$ km/t og $w = 141$ km/t.
- g) Finn sannsynligheten for type I feil dersom $w = 137$ km/t og vis resultatet i grafen i forrige punkt.

Oppgave 3 La $f(v|w)$ være den betingede sannsynlighetstettheten til en stokastisk variabel V gitt $W = w$ og la $f(w)$ være sannsynlighetstettheten til en stokastisk variabel W .

a) Vis at

$$E(V) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} v f(v|w) dv \right] f(w) dw \quad (2)$$

Farten W til biler på en vei med fartsgrense 90 km/t antas å være gammafordelt med formparameter α og skalaparameter β slik at $\mu = E(W) = \alpha\beta$ og $\sigma^2 = \text{Var}(W) = \alpha\beta^2$. Farten måles med en radar som spesifisert i Oppgave 2 og resultatet V er normalfordelt med middelerdi w og standardavvik γw når den reelle farten $W = w$. Det antas at $\gamma = 2\%$. Farten til 55 biler måles og rapporteres ved en empirisk middelerdi $\bar{v} = 89.2$ km/t og et empirisk standardavvik $s = 7.3$ km/t.

b) Vis at $E(V) = E(W) = \mu$.

c) Finn et forventningsrett estimat $\hat{\mu}$ for μ .

d) Beregn et estimat u av standardavviket[†] til estimatoren brukt i forrige punkt.

e) Beregn en verdi $U = ku$ slik at $\hat{\mu} \pm ku$ er et approksimativt 95% konfidensintervall for μ ved å bruke sentralgrenseteoremet og ved å approksimere standardavviket med estimatet u .

f) I følge det internasjonale standardiseringsorganet ISO skal en måling rapporteres med en standard usikkerhet u og en utvidet usikkerhet U som i forrige punkt. Gi en kort begrunnelse for dette ved å gi en tolkning av intervallet i forrige punkt.

g) Det kan vises at $\text{Var}(V) = (1 + \gamma^2)\sigma^2 + \gamma^2\mu^2$. Estimer σ , α og β .

[†]Da er u standard usikkerhet til estimatet $\hat{\mu}$ i følge ISO.

Oppgave 4 I denne oppgaven er U en stokastisk variabel som har uniform fordeling på intervallet $[0, 1)$.

- a) Definer hva det betyr at X er en stokastisk variabel og definer den kumulative fordelingen F_X til X .
- b) Skisser grafene til sannsynlighetstettheten f_U og den kumulative fordelingen F_U til U .
- c) La $A_1 = [1/2, 1)$ og $A_2 = [1/4, 1/2) \cup [3/4, 1)$. Vis at hendelsene $(U \in A_1)$ og $(U \in A_2)$ er uavhengige.
- d) Indikatorfunksjonen I_A til en mengde A er definert ved $I_A(x) = 1$ for $x \in A$ og $I_A(x) = 0$ for $x \notin A$. Vis at $X_1 = I_{A_1}(U)$ og $X_2 = I_{A_2}(U)$ er uavhengige Bernoulli fordelte stokastiske variabler.
- e) Indikatorvariablene $I_{B_1}, I_{B_2}, \dots, I_{B_n}$ er uavhengige dersom hendelsene i samlingen $\{B_1, B_2, \dots, B_n\}$ er uavhengige. Definer hva det betyr at hendelsene i samlingen $\{B_1, B_2, \dots, B_n\}$ er uavhengige. Konstruer uavhengige Bernoulli-fordelte stokastiske variabler $X_1, X_2, \dots, X_n$ hvor

$$X_i = I_{A_i}(U) \text{ og } P(X_i = 1) = 1/2 \quad (3)$$

Hint: Her kan A_1 og A_2 velges som gitt over og $A_3, A_4, \dots, A_n$ velges på en passende måte.[‡]

[‡]Fra resultatene i oppgave 4 følger det at $U = X_1/2 + X_2/2^2 + X_3/2^3 + \dots$ har uniform fordeling på intervallet $[0, 1)$ når $X_1, X_2, X_3, \dots$ er en følge av uavhengige Bernoulli fordelte stokastiske variabler med $P(X_i = 1) = 1/2$. Et tilfeldig tall U på intervallet $[0, 1)$ er dermed ekvivalent med en Bernoulli prosess tilsvarende en følge myntkast. Et desimaltall som 0.31415... tilsvarer en følge siffer 3, 1, 4, 1, 5, ... og en tilsvarende korrespondanse ved bruk av totalssystemet gir ideen i konstruksjonen i oppgaven.