

Institutt for matematiske fag

## Eksamensoppgave i **TMA4240 / TMA4245 Statistikk**

**Faglig kontakt under eksamen:** Håkon Tjelmeland

**Tlf:** 48 22 18 96

**Eksamensdato:** 10. august 2017

**Eksamenstid (fra–til):** 09.00-13.00

**Hjelpemiddelkode/Tillatte hjelpemidler:**

*Tabeller og formler i statistikk, Akademika,*

K. Rottmann: *Matematisk formelsamling,*

Kalkulator Casio fx-82ES PLUS, CITIZEN SR-270X, CITIZEN SR-270X College eller HP30S,

Gult stemplet A5-ark med egne håndskrevne notater.

**Annen informasjon:**

Alle svar skal begrunnes og besvarelsen skal inneholde naturlig mellomregning.

**Målform/språk:** bokmål

**Antall sider:** 3

**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    Sannsynlighet**

Anta at den stokastiske variabelen  $X$  har sannsynlighetstetthet

$$f(x; \theta) = c \exp\{-(x - \theta)\}, \quad x \geq \theta,$$

mens  $f(x; \theta) = 0$  for  $x < \theta$ .  $\theta$  betegner en konstant parameter og  $c$  er en positiv konstant.

- a) Bestem konstanten  $c$ , og finn sannsynligheten for at  $X > \theta + 1$ .
- b) Utled et uttrykk for sannsynlighetsmaksimeringsestimatoren  $\hat{\theta}$  for parameteren  $\theta$  basert på  $n$  uavhengige observasjoner,  $x_1, \dots, x_n$ , av  $X$ . Hint: Skissér rimelighetsfunksjonen (likelihood-funksjonen).
- c) La  $X_1, X_2, \dots, X_{10}$  betegne 10 uavhengige stokastiske variabler som hver har samme sannsynlighetstetthet som  $X$ . Bestem sannsynlighetstettheten for den stokastiske variabelen  $W = \min\{X_1, X_2, \dots, X_{10}\}$ . Regn ut sannsynligheten for at  $W > \theta + 1$ .

**Oppgave 2    Utmatting**

Rotorfestene på et helikopter er utsatt for brudd på grunn av utmatting. For å bedre utmattingsfastheten kan en foreta en overflateherding av festene ved å bruke en av de to metodene A og B beskrevet nedenfor.

Metode A, også kalt "shot-peening", går ut på at materialet utsettes for en type sandblåsing med ganske grove kuler.

Ved Metode B (nitring) blir overflaten varmet opp til ca. 500 grader celsius og tilført ammoniakk. Dette gir et tynt, slitesterkt sjikt.

En er interessert i å undersøke om metode A gir bedre utmattingsfasthet enn metode B. Ved fabrikken som produserer rotorfestene, velges 13 stykker tilfeldig til et forsøk. De 7 første rotorfestene behandles etter metode A, og de øvrige etter metode B. De 13 rotorfestene settes i en utmattingsrigg og utsettes for sykliske belastninger med spenningsvidde 700 MPa (MegaPascal). Erfaring har vist at den naturlige logaritmen til antall sykler før et brudd oppstår vil tilnærmet være en normalfordelt stokastisk variabel.

Den naturlige logaritmen til antall sykler før brudd for rotorfestene som er behandlet etter metode A, er  $X_1, \dots, X_7$ . For rotorfestene behandlet etter metode

B fås tilsvarende  $Y_1, \dots, Y_6$ . Alle de 13 observasjonene antas å være uavhengige, normalfordelte stokastiske variabler med ukjente forventningsverdier og varianser:

$$E(X_i) = \mu_A, \text{Var}(X_i) = \sigma_A^2, i = 1, \dots, 7.$$

$$E(Y_j) = \mu_B, \text{Var}(Y_j) = \sigma_B^2, j = 1, \dots, 6.$$

Anta i de to neste punktene at  $\mu_A = \mu_B = \mu$  og  $\sigma_A = \sigma_B = \sigma$ , men at deres felles verdier er ukjente.

a) Vis at

$$\hat{\mu} = \frac{1}{2}(\bar{X} + \bar{Y})$$

er en forventningsrett estimator for  $\mu$  og finn dens varians uttrykt ved  $\sigma$ . Her er  $\bar{X} = \sum_{i=1}^7 X_i/7$  og  $\bar{Y} = \sum_{j=1}^6 Y_j/6$ .

b) Foreslå så en forventningsrett estimator  $\mu^*$  for  $\mu$  som har mindre varians enn  $\hat{\mu}$ . Vis dette ved utregning.

I resten av oppgaven vil vi ikke lenger anta at  $\mu_A = \mu_B$  og  $\sigma_A = \sigma_B$ . Følgende formel er hentet direkte fra læreboken:

$$\nu = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$

Den brukes til å beregne tilnærmet antall frihetsgrader  $\nu$  for estimatorer som er vanlig å bruke ved estimering og hypotesetesting som i resten av denne oppgaven.

c) La  $\delta = \mu_A - \mu_B$ . Utled et  $100(1 - \alpha)\%$  konfidensintervall for  $\delta$ . Hva blir intervallet når  $\alpha = 0.1$  med de empiriske resultatene som er gitt til slutt i oppgaven?

d) Tyder resultatene fra forsøket på at metode A gir bedre utmattingsfasthet enn metode B? Formulér dette som et hypotesetestingsproblem og begrunn valget av nullhypotese og alternativ hypotese. Velg signifikansnivå 0.05 og gjennomfør testen med de resultatene som er gitt til slutt i oppgaven.

$$\bar{x} = 15.22$$

$$\sum_{i=1}^7 (x_i - \bar{x})^2 = 0.32$$

$$\bar{y} = 14.56$$

$$\sum_{j=1}^6 (y_j - \bar{y})^2 = 0.47$$

$$\nu \approx 9$$

### Oppgave 3      Kvalitetskontroll

Ved produksjonen av en type medisin i pulverform, som selges på forseglede bokser pakket i pappesker med  $k$  bokser i hver ekse, er det viktig å sørge for at pulveret ikke er forurenset av soppsporer. Dette kontrolleres ved en laboratorietest. Sannsynligheten for at en tilfeldig valgt boks gir et positivt testresultat (dvs. at boksen er infisert), antas å være lik  $p$ , og testresultatene for ulike bokser antas å være uavhengige.

Man går fram på følgende måte: Tester gjøres av de  $k$  ( $> 1$ ) boksene i en eske om gangen. Vi kaller disse boksene en  $k$ -gruppe. Halvparten av innholdet i hver boks i en  $k$ -gruppe blandes og blandingen analyseres. Hvis den gir positivt testresultat, noe som betyr at minst en av boksene er infisert, gjøres nye tester av resten av innholdet i de  $k$  boksene, og prøvene analyseres enkeltvis. Hvis derimot blandingen gir negativt testresultat, antas at ingen av boksene i  $k$ -gruppen er infisert. Hvis det er soppsporer i en boks, kan de antas tilnærmet jevnt fordelt i boksen.

- a) Begrunn hvorfor antall infiserte bokser i en  $k$ -gruppe er binomisk fordelt, og vis at sannsynligheten for at en blanding av innholdet i  $k$  bokser skal gi positivt testresultat, er

$$1 - (1 - p)^k.$$

- b) Hvordan defineres betinget sannsynlighet for en hendelse  $A$  gitt en annen hendelse  $B$ ? Hvis en boks er i en  $k$ -gruppe som har vist positivt testresultat, hva er da sannsynligheten for at boksen er infisert?
- c) Anta at en kvalitetskontroll gjennomføres ved at  $m$  esker med  $k$  bokser i hver tas ut for å testes. La  $X$  betegne antall prøver som må analyseres før hele partiet på  $mk$  bokser er ferdig testet. Vis at

$$E(X) = m + mk \left( 1 - (1 - p)^k \right).$$

Hvis  $k = 4$ , for hvilke verdier av  $p$  er den benyttede framgangsmåten å foretrekke i stedet for å teste boksene enkeltvis med en gang?