NTNU
Kunnskap for en bedre verden

Institutt for matematiske fag

Eksamensoppgave i **TMA4295 Statistisk Inferens**

Faglig kontakt under eksamen: Professor Henning Omre
Tlf: 90937848

Eksamensdato: November 29.2016
Eksamenstid (fra–til): 09:00-13:00
Hjelpemiddelkode/Tillatte hjelpemidler: C.
Tabeller og Formler i Statistikk, Tapir
NTNU-godkjent kalkulator
Personlig, håndskrevet, gult huskeark - A5-format

Målform/språk: bokmål
Antall sider: 5
Antall sider vedlegg: 6

Kontrollert av:

Informasjon om trykking av eksamensoppgave	
Originalen er:	
1-sidig ☐	**2-sidig** ☒
sort/hvit ☒	**farger** ☐
skal ha flervalgskjema ☐	

Dato Sign

Oppgave 1 Buffons Nål

I 1773, designet Georges-Louis Leclerc (1707-1788) - titulert 'Compte de Buffon' - et forsøk og definerte en estimator for å beregne verdien av den fysiske konstanten π - kjent for oss idag som 3.1415926535... 'The Compte' estimerte π, men det eksisterte ingen statistisk teori på den tid så han hadde ingen dyp innsikt i sin estimeringsprosedyre. En utvidelse av det originale forsøket ble også utført av 'the Compte' i 1777 - men uheldigvis var det en feil i hans utregninger - og den korrekte løsning ble presentert av Pierre-Simon Laplace (1749-1827) i 1812. Gjennom historien etter denne tid, har Buffons nål-problem vært omfattet med stor interesse fra statistikere. Vi skal re-beregne noen av disse resultatene i denne oppgaven.

Betrakt Buffons nål-forsøk som skissert i Figur 1a. I 2D planet defineres det et sett av parallelle linjer med innbyrdes avstand $d > 0$. En nål med lengde $0 > l \leq d$ blir kastet på 2D planet med uniformt fordelt lokasjon og orientering. Vi observerer den tilfeldige variablen X:

$$X \rightarrow f(x|p_0, p_1) = \begin{cases} 0 & \text{- ingen kryssing nål og linje} \quad \text{med sannsynlighet } p_0 \\ 1 & \text{- en kryssing nål og linje} \quad \text{med sannsynlighet } p_1 \end{cases}$$

hvor sannsynlighetene (p_0, p_1) kan regnes ut fra geometrien i forsøksdesignet:

$$p_0 = 1 - 2\frac{l}{d}\frac{1}{\pi} = 1 - 2r\frac{1}{\pi}$$
$$p_1 = 2\frac{l}{d}\frac{1}{\pi} = 2r\frac{1}{\pi}$$

med $r = l/d$. Merk at det tilfeldige forsøket er et Bernoulli forsøk med parameter $p = p_1$.

Utfør forsøket n ganger og observer det tilfeldige utvalget:

$$\mathbf{X}_n : X_1, \ldots, X_n \text{ iid } f(x|p_0, p_1)$$

Fokus for forsøket er på beregning av konstanten π, men først skal parametren $p = 2r/\pi$ evalueres.

a) Utled et uttrykk for en suffisient observator for p basert på $\mathbf{X}_n$.

Er denne suffisiente observatoren også komplett ? Begrunn svaret.

Demonstrer at uttrykket for maximum likelihood (ML) estimatoren for p basert på $\mathbf{X}_n$ er,

$$\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

b) Er ML estimatoren $\hat{p}_n$ forventningsrett for p ? Begrunn svaret.

Er ML estimatoren $\hat{p}_n$ også uniformt minimum varians forventningsrett (UMVU) estimatoren for p ? Begrunn svaret.

Betrakt forsøksdesignet - hvordan kan det bli satt opp slik at en oppnår minst mulig varians i estimatoren $\hat{p}_n$?

Fokuser nå på konstanten π.

c) Spesifiser uttrykket for ML estimatoren $\hat{\pi}_n$ for π basert på $\mathbf{X}_n$. Begrunn svaret.

Utled uttrykket for den asymptotiske variansen for estimatoren $\hat{\pi}_n$.

Er estimatoren $\hat{\pi}_n$ en asymptotisk effisient estimator for π ? Begrunn svaret.

d) Utled uttrykket for forkastningsområdet for den asymptotiske likelihoodforholds (LR) α-nivå testen for hypotesen

$$\mathcal{H}_0 : \pi = \pi_0 = 3.14 \text{ versus } \mathcal{H}_1 : \pi \neq \pi_0$$

Spesifiser uttrykket for det tilhørende asymptotiske LR $(1-\alpha)$-konfidensområdet for π. Begrunn svaret.

'The Compte' utvidet forsøket i 1777 som skissert i Figur 1b, og dette forsøket ble senere omtalt som Buffon-Laplaces nål-forsøk. I 2D planet defineres to ortogonale sett med parallelle linjer, med samme innbyrdes avstand $d > 0$. En nål med lengde $0 < l \leq d$ blir kastet på 2D planet med uniformt fordelt lokasjon og orientering. Observer den tilfeldige variablen X ,

$$X \to f(x|p_0, p_1, p_2) = \begin{cases} 0 & \text{- ingen kryssing nål og linje} \quad \text{med sannsynlighet } p_0 \\ 1 & \text{- en kryssing nål og linje} \quad \text{med sannsynlighet } p_1 \\ 2 & \text{- to kryssinger nål og linjer} \quad \text{med sannsynlighet } p_2 \end{cases}$$

hvor sannsynlighetene (p_0, p_1, p_2) kan regnes ut fra geometrien i forsøket,

$$p_0 = 1 - (4 - r)r\frac{1}{\pi}$$
$$p_1 = 2(2 - r)r\frac{1}{\pi}$$
$$p_2 = r^2\frac{1}{\pi}$$

med $r = l/d$.

Utfør forsøket n ganger og observer det tilfeldige utvalget:

$$\mathbf{X}_n : X_1, \ldots, X_n \text{ iid } f(x|p_0, p_1, p_2)$$

Fokus for det utvidete forsøket er også på estimering av konstanten π, men først skal $\theta = 1/\pi$ evalueres og deretter skal π evalueres.

e) Demonstrer at ML estimatoren $\tilde{\theta}_n$ for θ basert på $\mathbf{X}_n$ er

$$\tilde{\theta}_n = \frac{1}{n(4-r)r}[N_1 + N_2]$$

med $N_j = \Sigma_{i=1}^n I(X_i = j)$ for $j = 1, 2$ hvor indikatorfunksjonen $I(A)$ er 1 hvis A er sann og 0 ellers.

Demonstrer at denne ML estimatoren $\tilde{\theta}_n$ også er UMVU estimatoren for θ.

f) Spesifiser ML estimatoren $\tilde{\pi}_n$ for π. Begrunn svaret.

Utled uttrykket for den asymptotiske variansen til estimatoren $\tilde{\pi}_n$.

Utled uttrykket for den asymptotiske relative effisiensen (ARE) til $\tilde{\pi}_n$ med hensyn til $\hat{\pi}_n$ fra Punkt c.

Sett inn $l = d$, dvs $r = 1$, samt $\pi = 3.14$ i uttrykket for ARE og beregn dens numeriske verdi. Tolk resultatet.

Oppgave 2

Bruk

$$X \to f(x|\theta) = Uni[\theta - 1/2, \theta + 1/2] \; ; \; x \in \mathcal{R} \; ; \; \theta \in \mathcal{R}$$

som medfører at den tilfeldige variablen X er uniformt fordelt på intervallet $[\theta - 1/2, \theta + 1/2]$ med ukjent modellparameter θ.

Betrakt det tilfeldige utvalget,

$$\mathbf{X}_n : X_1, \ldots, X_n \text{ iid } f(x|\theta)$$

Det tilhørende økende ordnede tilfeldige utvalget er benevnt $X_{(1)}, \ldots, X_{(n)}$.

Definer en tilfeldig variabel $U \to Uni[0,1]$ med tilfeldig utvalg $\mathbf{U}_n : U_1, \ldots, U_n$ med tilhørende økende ordnet tilfeldig utvalg $U_{(1)}, \ldots, U_{(n)}$. De tilhørende sannsynlighetstetthetene er,

$$U_{(i)} \to f_{U_{(i)}}(u_i) = \frac{n!}{(i-1)!(n-i)!} u_i^{i-1}(1-u_i)^{n-i} \; ; \; 0 \le u_i \le 1$$

det vil si Beta fordelt $Beta(i, n-i+1)$ (se Tabeller og Formler i Statistikk),

$$[U_{(i)}, U_{(j)}] \to f_{U_{(i)}U_{(j)}}(u_i, u_j) = \frac{n!}{(i-1)!(j-1-i)!(n-j)!} u_i^{i-1}(u_j - u_i)^{j-1-i}(1-u_j)^{n-j}$$
$$; \; 0 \le u_i \le u_j \le 1$$

Definer estimatoren for θ,

$$\hat{\theta}_n = \frac{1}{2}[X_{(1)} + X_{(n)}]$$

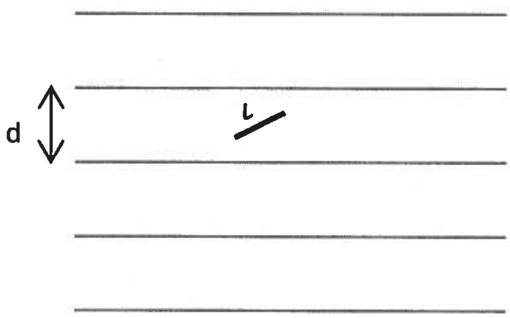
a) Utled uttrykk for de suffisiente observatorene for θ basert på utvalget $\mathbf{X}_n$, ved å bruke 'the Factorization theorem'.

Demonstrer at estimatoren $\hat{\theta}_n$ er maksimum likelihood (ML) estimatoren såvel som en forventningsrett estimator for θ.

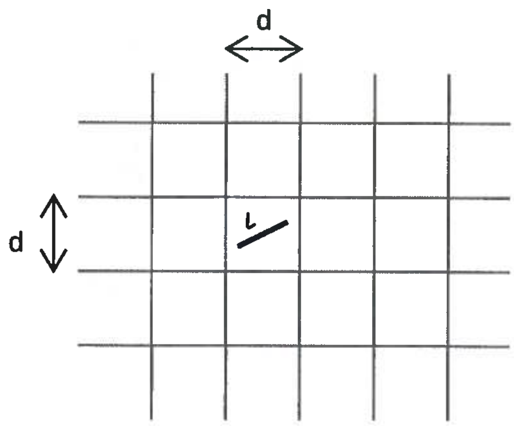
b) Utled uttrykket for det eksakte $(1-\alpha)$-konfidensintervallet for θ basert på estimatoren $\hat{\theta}_n$.

Georges-Louis Leclerc, Comte de Buffon

Pierre-Simon Laplace



(a) Buffons nål



(b) Buffon-Laplaces nål

Figur 1: Forsøksdesign